INVESTIGACION Y
CIENCIA

Edición en español de SCIENTIFIC AMERICAN



AVIONES ULTRALIGEROS

Septiembre 1982
300 PTAS.

Los espacios en gris
corresponden a publicidad
en la edición impresa

Número 72
Septiembre 1982

INVESTIGACION Y
CIENCIA *Edición en español de* SCIENTIFIC AMERICAN

Colaboradores de este número:

Asesoramiento y traducción:

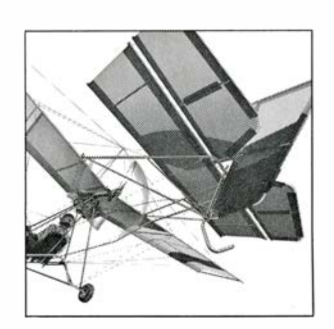
Manuel Puigcerver: *El Telescopio Espacial* y *Efusiones de energía en estrellas jóvenes;* Manuel Ballester Boix: *Superconductores orgánicos;* Juan José Martínez García: *Aviones ultraligeros;* Quira Cadenas y Eduardo Cadenas: *Topoisomerasas de ADN;* Joandomènec Ros: *La reproducción cooperativa de la abubilla arbórea verde;* Angel Castañer: *Aneurismas;* Luis Bou: *Temas metamágicos;* J. Vilardell: *Taller y laboratorio.*

Ciencia y sociedad:

Ramón Margalef

Libros:

José Obiols Salvat, José Luis Comellas, Luis Mas y Miguel Alcaraz



LA PORTADA

La ilustración de la portada muestra un avión ultraligero, el Quicksilver MX, de diseño mucho más avanzado que el primer vehículo, construido hace siete años montando un motor de Kart de 12 HP en un planeador biplano de piloto suspendido (véase "Aviones ultraligeros", por Michael A. Markowski, en este mismo número). El Quicksilver MX, fabricado por Eipper-Formance Inc., tiene un motor Cuyuna de 30 HP, diseñado especialmente para vuelo ultraligero. El vehículo despega y aterriza a unos 32 kilómetros por hora, asciende a 45 kilómetros por hora y tiene una velocidad de crucero de 70 kilómetros por hora. Las siglas MX (de multieje –"multiaxes"– en inglés) aluden al sistema de control, que actúa sobre tres ejes independientes: el piloto maneja los controles que mueven el timón de dirección y las superficies de ala y cola para variar el asiento, la guiñada y el balance. Esta palanca se maneja con la mano derecha. El mando de control visible en la ilustración es la palanca de gases, que se maneja con la mano izquierda. Tiene una envergadura de 9,7 metros y pesa 90 kilos.

Suscripciones:

Prensa Científica, S. A.
Calabria, 235-239
Barcelona-29 (España)
Teléfono 322 05 51 ext. 37

Condiciones de suscripción:

España:
Un año (12 números): 3.300 pesetas
Extranjero:
Un año (12 números): 52 U.S. $
Ejemplar atrasado ordinario:
340 pesetas
Ejemplar atrasado extraordinario:
440 pesetas

Distribución para España

Distribuciones de Enlace, S. A.
Ausias March, 49, Barcelona-10

Distribución para los restantes países:

Editorial Labor, S. A.
Calabria, 235-239 - Barcelona-29

Publicidad:

Madrid:
Gustavo Martínez Ovin
Avda. de Moratalaz, 137, Madrid-30
Tel. 430 84 81
Cataluña:
Lourdes Padrós
Manuel Girona, 61, 2.°, Barcelona-34
Teléfono 204 45 83

Controlado
por O.J.D.



PROCEDENCIA DE LAS ILUSTRACIONES

Pintura de la portada de Ted Lodigensky

Página	Fuente
7-8	Perkin-Elmer Corporation
9-15	George V. Kelvin
16-17	John L. Tonry, del Instituto de Estudios Avanzados
21-27	Gabor Kiss
28	Cortesía del Centre National de la Recherche Scientifique (CNRS)
30	Gabor Kiss
33-37	Walken Graphics
38	Michael A. Markowski
45	David F. Malin, Observatorio Anglo-australiano y Photolabs, Royal Observatory, Edimburgo
46	© National Geographic Society-Palomar Observatory Sky Survey
47	J. T. Williams, Multiple Mirror Telescope Observatory
48-50	Ilil Arbel
51	© National Geographic Society-Palomar Observatory Sky Survey (*izquierda*), Ilil Arbel (*derecha*)
52	© National Geographic Society-Palomar Observatory Sky Survey
53-54	Ilil Arbel

Página	Fuente
57	James C. Wang, Universidad de Harvard
58	Ian Worpole
59	Ian Worpole (*arriba*); James C. Wang, Universidad de Harvard (*abajo*)
60-68	Ian Worpole
71	Miguel Alonso
72-82	A. Ilundáin, R. J. Naftalin y M. Alonso
85-87	Patricia J. Wynne
88-92	Ilil Arbel
95-99	Carol Donner
100	Kaj Johansen, Universidad de Washington (*izquierda*); Howard Ricketts, Facultad de Medicina de la Universidad de Washington (*derecha*)
101	Howard Ricketts, Universidad de Washington (*arriba izquierda y abajo izquierda*); Carol Donner (*arriba derecha*); Laurence D. Cromwell, Universidad de Washington (*abajo derecha*)
102	Kaj Johansen, Facultad de Medicina de la Universidad de Washington
103-114	Jerome Kuhl
117	R. F. Bonifield (*arriba*), William M. Strouse (*abajo*)
118-121	Michael Goodman

ISSN 0210-136X
Dep. legal: B. 38.999-76
Fotocomposición Tecfa
Pedro IV, 160 - Barcelona-5
Fotocromos reproducidos por GINSA, S.A.
Imprime GRAFESA
Gráfica Elzeviriana, S. A.
Nápoles, 249 - Tel. 207 40 11
Barcelona-13
Printed in Spain - Impreso en España

Los autores

JOHN N. BAHCALL y LYMAN SPITZER, JR., ("El Telescopio Espacial") son astrofísicos que han trabajado desde su primera fase en el proyecto que da título al artículo. Bahcall es profesor del Instituto de Estudios Superiores y encargado de curso de astronomía en la Universidad de Princeton. Cursó el primer ciclo de estudios en la Universidad de California en Berkeley, y el segundo, que completó en 1957, en la de Chicago. En 1961 se recibió de doctor en física por la Universidad de Harvard. De 1960 a 1962 disfrutó de una beca en la Universidad de Indiana. Ese último año se incorporó al Instituto de Tecnología de California, donde desempeñó tareas de investigación y docentes. En 1968, ingresó en el cuerpo docente del Instituto de Estudios Superiores. Spitzer acaba de dejar, por jubilación, la cátedra de astronomía Charles A. Young de la Universidad de Princeton. Licenciado por la Universidad de Yale en 1935, se doctoró por la de Princeton en 1938. De 1939 a 1942 enseñó en la de Yale. Durante la Segunda Guerra Mundial colaboró en varios proyectos militares; entre otros, uno sobre el sonar. En 1946 regresó a Yale como profesor asociado. En 1947 se trasladó a Princeton como docente y director del departamento de ciencias astrofísicas y director del observatorio de la Universidad. Conservó ambos cargos hasta 1979. Fue uno de los primeros defensores de la plataforma espacial para observaciones astronómicas.

KLAUS BECHGAARD y DENIS JEROME ("Superconductores orgánicos"), quimicofísico y físico, respectivamente, vienen trabajando desde 1975 en el tema que desarrollan en esta colaboración. Bechgaard enseña química en la Universidad de Copenhague, por la que se doctoró en 1973. Tras un año en la Johns Hopkins volvió a Copenhague, donde inició sus investigaciones sobre superconductores orgánicos. Se encuentra en este momento con permiso sabático en la Universidad Paris-Sud en Orsay. Jérome pertenece al Centre National de Recherche Scientifique (CNRS) y trabaja en el laboratorio que este organismo tiene en la Universidad Paris-Sud. En 1965 se doctoró por esta institución superior. En 1967 se incorporó al laboratorio del CNRS, después de haber realizado un año de estudios postdoctorales en la Universidad de California en San Diego. Jérome investiga también el magnetismo y las transiciones entre fases de la materia.

MICHAEL A. MARKOWSKI ("Aviones ultraligeros") es autor de artículos sobre aeronaves pequeñas y ligeras y asesor de compañías que las fabrican. Se formó en la Universidad estatal de Pennsylvania. Trabajó luego como ingeniero aeronáutico: en 1968 estuvo en la McDonnell Douglas Corporation; de 1969 y 1971 trabajó para la empresa Sikorsky Aircraft y, de 1973 a 1975, para la Atkins & Merrill Inc., compañía especializada en el ensayo de diseños de avión. Ha sido también socio de una compañía de fabricación de planeadores de piloto suspendido. Desde el año 1975, el interés principal de Markowski se centra en la aviación ultraligera.

CHARLES J. LADA ("Efusiones de energía en estrellas jóvenes") enseña astronomía en la Universidad de Arizona. Inició su carrera en Boston, completándola en Harvard, donde se doctoró en 1975. El bienio de 1975 a 1977 lo pasó en el Centro de Astrofísica del Observatorio del Harvard College y del Observatorio Astrofísico Smithsoniano. En 1977 y 1978 fue miembro investigador de Harvard. En 1978 se trasladó a Arizona, donde ejerció funciones docentes y trabajó al propio tiempo en el Observatorio Steward. Acaba de recibir el premio Bart J. Bok de investigación astrofísica, en su edición de 1982, que la Universidad de Harvard concede con periodicidad anual.

JAMES C. WANG ("Topoisomerasas de ADN") es profesor de bioquímica y biología molecular de la Universidad de Harvard. Nació en China continental; en 1949 se trasladó, con su familia, a Taiwan. Se licenció en ciencias por la Universidad Nacional de Taiwan, en 1959, en la que enseñó en 1959 y 1960. Ese año se trasladó a los Estados Unidos para ampliar estudios. Se licenció por la Universidad de Dakota del Sur y se doctoró por la de Missouri. De 1964 a 1966 investigó en el Instituto de Tecnología de California. En 1966 se trasladó a la Universidad de California en Berkeley, en cuyo claustro docente ingresó en 1974. Tres años después se trasladaba a Harvard.

ANUNCIACION ILUNDAIN y R. J. NAFTALIN ("La secreción intestinal") han trabajado juntos asiduamente en el laboratorio del segundo en el King's College de la Universidad de Londres. Ilundáin procede de la Universidad de Navarra y pertenece al cuerpo de profesores adjuntos por oposición. Desde 1980 se halla adscrita al Departamento Interfacultativo de Fisiología de la Universidad de Navarra. Naftalin estudió en la Universidad de Glasgow. Desde 1968 hasta 1975 enseñó fisiología en la Universidad de Leicester. Pasó un año (1972) en la de Yale. Su principal tema de investigación versa sobre el transporte de azúcares y electrolitos a través del epitelio intestinal.

J. DAVID LIGON y SANDRA H. LIGON ("La reproducción cooperativa de la abubilla arbórea verde") son biólogos que de forma intermitente, y desde 1975, han venido estudiando conjuntamente las abubillas arbóreas verdes en Kenia. David Ligon es profesor de biología de la Universidad de Nuevo México. Recibió su formación en las universidades de Oklahoma, Florida y Michigan; por esta última se doctoró en 1967. En 1967 y 1968 enseñó biología en la Universidad estatal de Idaho. Pasó a la Universidad de Nuevo México en 1968, de la que es profesor desde 1977. Su principal campo de investigación se refiere a las ventajas adaptativas que el comportamiento social confiere a las aves. Sandra Ligon profesa biología en la Universidad de Nuevo México. Cursó el primer ciclo de ciencias en el Knox College, licenciándose por la Universidad de Nuevo México en 1973. A continuación trabajó durante dos años en el National Fish and Wildlife Laboratory de los Estados Unidos. Sus investigaciones se centraron especialmente en los sirenios, orden de mamíferos acuáticos herbívoros donde se integran los manatíes y los dugongos.

KAJ JOHANSEN ("Aneurismas") enseña cirugía en la Universidad de Washington. Allí cursó sus estudios superiores. Después de realizar su residencia en cirugía, se doctoró en fisiología y biofísica por la Universidad de California en San Diego, en 1977. En 1978 volvió a la Universidad de Washington, donde dirige hoy el programa de residentes de cirugía y da clases en la Facultad de Medicina. Cuenta Johansen: "Soy un gran aficionado a carreras de larga distancia, y he corrido aproximadamente 35 maratones. Fui campeón norteamericano de la carrera de 50 kilómetros en 1974, con un tiempo de 2 horas 57 minutos 00 segundos; marca récord hasta 1979.»

Hace...

José M.ª López Piñero

... cuatrocientos años

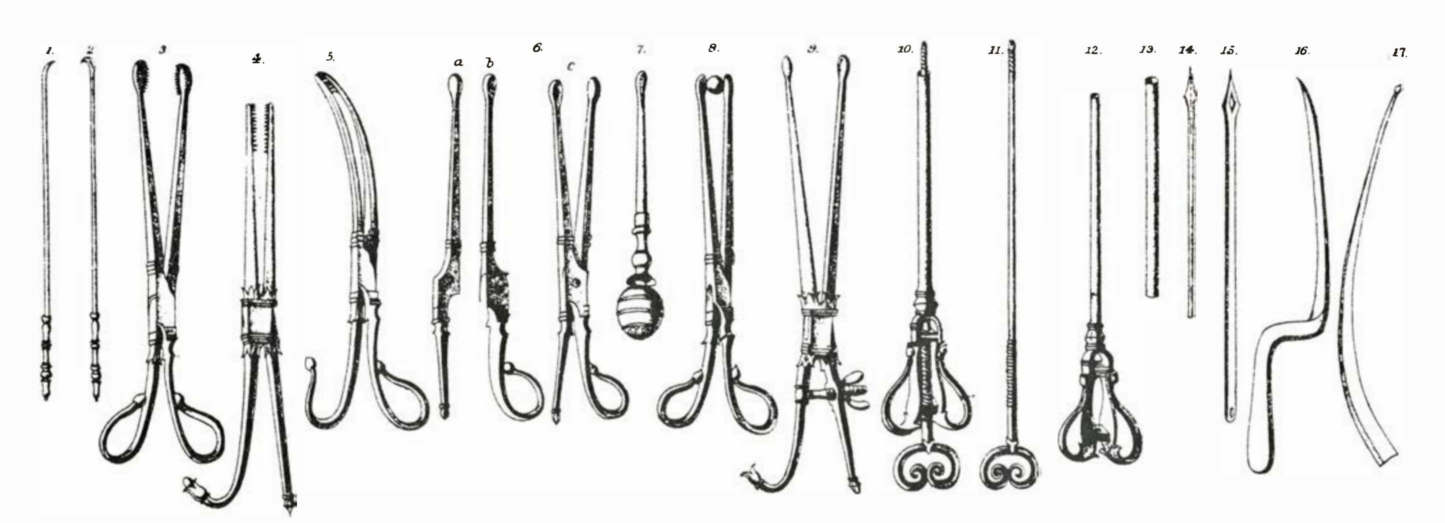
Dionisio Daza Chacón publicó la primera parte de su *Practica y Theorica de Cirugia*, uno de los principales tratados quirúrgicos aparecidos en la Europa del siglo XVI.

Hijo de un médico vallisoletano y nieto de Dionisio Chacón, médico de cámara de Felipe I, Daza nació en Valladolid el año 1510. Estudió en la Universidad de Valladolid y más tarde en la de Salamanca, donde obtuvo el título de licenciado en medicina. Consagró toda su vida a la cirugía, siendo durante largos años la gran figura quirúrgica de los ejércitos de Carlos I y más tarde de Felipe II, en los que tuvo como colega a Andrés Vesalio, el famoso renovador del saber anatómico, al que le unió una firme amistad. Ambos asistieron también a varias célebres consultas, entre ellas la motivada por las heridas que una caída le había producido en la cabeza al príncipe Don Carlos, hijo de Felipe II (1562). Participó en las grandes campañas de los ejércitos españoles en Flandes, Alemania y el Mediterráneo, y asistió a la batalla de Lepanto. Durante los períodos intermedios ejerció en la Corte y en sus hospitales. Fue nombrado cirujano del Hospital Real de Valladolid en 1561 y cirujano de cámara de Don Juan de Austria once años más tarde. En 1580, "atento a su edad y a los años que ha que me sirve... y que está muy enfermo", Felipe II le concedió la jubilación con todo su salario. Los años que le restaron de vida, hasta su fallecimiento en 1596, los repartió entre Madrid y Valladolid. Fue entonces cuando redactó el tratado que acabamos de citar, a cuya primera parte siguió una segunda que apareció originalmente en 1595. La obra se reeditó después en varias ocasiones.

Aparte de contener abundantes contribuciones originales, el tratado de Daza Chacón es una exposición sistemática de los saberes y las técnicas quirúrgicas, valorados críticamente a través de una continua referencia a su experiencia personal. Ofrece una crónica de las contiendas de la época desde la perspectiva del cirujano tan viva como los *Voyages* del francés Ambroise Paré. Está redactado en un castellano castizo y directo, porque sus destinatarios eran los cirujanos propiamente dichos, que en esta época carecían de formación universitaria y desconocían el latín. Daza justifica el uso de la lengua vulgar en los siguientes términos: "Te quiero decir la ocasión que me movió a escribir en nuestra lengua española antes que en latín... habiendo yo examinado en la Corte más de doce años a los cirujanos romancistas que se iban a examinar, vi muchos que tenían muy buenas habilidades, que por falta de tener libros en su lengua estaban muy atrás de lo que pudieran saber, determiné darles todo lo bueno que los antiguos y modernos dijeron acerca desta materia".

Por otra parte, Daza defiende la importancia de la práctica en una línea abiertamente opuesta al desprecio del trabajo manual propio del pensamiento tradicional y, al mismo tiempo, tiene clara conciencia de la superioridad del cirujano con estudios universitarios frente a los meros empíricos, que ridiculiza en numerosas ocasiones. Afirma arrogantemente que "el buen cirujano es el verdadero médico". Su superioridad reside precisamente en el carácter práctico de su ocupación: "La cirugía práctica, o agente, o factiva, como la quisiéredes llamar, es la verdadera cirugía, porque es saber poner por obra y ejercitar y hacer con las manos y con los instrumentos lo que el otro no supo muy bien parlar. Y ésta se pone entre las artes mecánicas, y no se alcanza ni puede uno ser perfecto en ella si no la hubiere ejercitado muchos años".

Aunque la formación originaria de Daza fuera de tipo bajomedieval, la convivencia con figuras como Andrés Laguna y Vesalio lo convirtieron en un hombre abierto a las nuevas corrientes. Por ello, se opuso a los criterios del galenismo arabizado en múltiples cuestiones y complementó con datos de autopsia varias de sus excelentes historias clínicas. Anotemos, como ejemplo, que una de ellas corresponde a la descripción de lo que hoy llamamos "blast injury": "Vi a dos en la guerra que al uno le pasó por entre las piernas una bala de una culebrina, y llevóle solamente los tafetanes de las calzas, y sin hacerle herida ninguna murió dentro de doce horas, con hacerle todos los beneficios posibles: hinchósele el muslo y inflamósele bravamente. Y sola la causa fue que, de la gran furia que llevaba la pelota, no sólo el aire penetró en el muslo pero quebróle y desmenuzóle los huesos dél; lo cual vimos haciendo disección dél".



Instrumentos quirúrgicos utilizados por Daza Chacón para el tratamiento de las heridas por arma de fuego

Las aportaciones de Daza corresponden a los más variados campos quirúrgicos. Ideó, por ejemplo, una técnica para el tratamiento operatorio de los aneurismas, ligando el vaso por encima del saco e incidiéndole después, y otra para la amputación, con desplazamiento del colgajo cutáneo sobre el muñón. También son muy notables los estudios que dedica a los tumores malignos y a la trepanación craneal. En este último punto tiene un criterio intervencionista, frente a los que mantenían una postura opuesta a las operaciones craneales. Con tal motivo, refleja la conciencia de superioridad que tenían los cirujanos españoles e italianos de la época respecto de los del resto de Europa, reduciendo la práctica de los cirujanos alemanes en los casos de heridas craneales a algo ridículo: "Estando yo en Alemania, donde estuve muchos años al servicio del invictísimo emperador Carlos V, nuestro señor de gloriosa memoria, vi que los cirujanos de aquella tierra, no sólo las heridas simples de cabeza, pero las compuestas y con grandes fracturas, las curaban con solas bebidas, a las cuales llamaban pigmenta, y en las heridas ponían unas hojas de berza, y los heridos comían y bebían a su gusto... Si alguno acaecía a sanar (que eran pocos) eran por ser robustos y fuertes, y tener grandísima y buena complexión. Pero, dejando todo lo dicho aparte como cosa de risa, la verdadera manera de curar las fracturas (craneales) es que por fuerza hemos de venir a echar mano de las herramientas".



Retrato de Daza Chacón que figura en la edición de 1582 de la primera parte de su **Practica y Theorica de Cirugia**

Destaca también por su particular riqueza la exposición que hace de las heridas, tema sobre el que había acumulado una experiencia excepcional durante sus años de cirujano militar. A este respecto, nos limitaremos a recordar que Daza Chacón fue uno de los primeros cirujanos europeos que adoptó el tratamiento "suave" de las heridas por arma de fuego, en contra de la doctrina arabizada del "pus loable" y tras desmentir su naturaleza venenosa y combusta. En su tratado, expone las circunstancias en las que Vesalio y él recibieron en plena campaña la noticia del nuevo procedimiento por parte del propio Bartolomeo Maggi, así como la decisiva influencia que la posterior recomendación de Andrés Laguna tuvo para que ambos lo adoptaran: "Como tenían por muy cierto que estas heridas (por arma de fuego) eran venenosas, entraban cauterizando y embutían las heridas de lechinos muy empapados en trementina y aceite de saúco muy hirviendo; y de esta manera de curar usamos el año de cuarenta y cuatro, estando la Majestad del emperador Carlos Quinto, de gloriosa memoria, sobre Landresi, y esta usaba el doctísimo Vesalio; con la cual manera de cura, no sólo los heridos eran infestados con grandísimos dolores y otros accidentes perniciosos, pero las llagas se hacían consórdidas y pútridas, que no nos podíamos averiguar con ellas. Estando luego adelante la Majestad Católica con su ejército sobre Sandesier, vino al campo un cirujano italiano que se llamaba Micer Bartolomé, muy docto y de mucha experiencia, y comenzó a curar estas heridas muy de otra manera que nosotros las curábamos, que era como si curara una herida contusa... que, como no martirizaba los heridos como nosotros hacíamos con los cauterios, todo le sucedía bien y curaba en brevísimo tiempo, respecto de lo que a nosotros nos duraban las curas. Visto y entendido el negocio (por los buenos sucesos), determinamos de seguir su parecer, y con el tuvimos muchos y buenos sucesos, a lo menos quedando yo en aquella fuerza por mandado de Su Majestad con quinientos heridos por nomina, sanaron tantos dellos... y todos se curaron de la manera que curaba el italiano; y tengo por cierto que, si se curaran de la otra, perecieran muchos. Acertó también en aquella sazón a venir al ejército el Doctor Laguna, el que comentó a Dioscórides, y nos aprobó la cura de Micer Bartolomé".

El Telescopio Espacial

En 1985, el transbordador espacial norteamericano pondrá en órbita el mayor telescopio astronómico proyectado para operar fuera de los efectos perturbadores de la atmósfera terrestre

John N. Bahcall y Lyman Spitzer, Jr.

La atmósfera terrestre es una ventana imperfecta sobre el universo. Las ondas electromagnéticas de la parte óptica del espectro (ondas más largas que los rayos X y más cortas que las de radio) penetran hasta la superficie de la Tierra sólo a través de unas pocas y estrechas bandas espectrales. La más ancha de las bandas transmitidas viene a corresponder a los colores de la luz visible; las ondas de las regiones ultravioleta e infrarroja vecinas las absorbe, casi en su integridad, la atmósfera. Por si fuera poco, la turbulencia que caracteriza a ésta desfigura las imágenes de los objetos celestes, aun cuando se contemplen con los más potentes telescopios instalados en nuestros observatorios terrestres.

Viene de lejos, pues, la ponderación de las ventajas que resultarían si se acometieran observaciones astronómicas desde fuera de la atmósfera. En los últimos decenios se ha adquirido bastante pericia en el manejo remoto de telescopios transportados por encima de la mayor parte, o incluso de la totalidad, de la atmósfera a bordo de cohetes suborbitales, globos para grandes alturas y satélites artificiales. Estas empresas han aportado descubrimientos importantes y obligado a revisar algunas teorías sobre la estructura y evolución del universo.

El paso siguiente a dar en este programa de exploración lo constituye el Telescopio Espacial, que pondrá en órbita alrededor de la Tierra el transbordador espacial de los Estados Unidos en 1985, de cumplirse lo proyectado. El ingenio será un telescopio reflector corriente con posibilidades poco corrientes. Será el mayor telescopio astronómico jamás puesto en órbita. Y la primera instalación científica internacional a largo plazo situada en el espacio.

De acuerdo con el proyecto, el Telescopio Espacial, en fase de construcción, constituirá un observatorio astronómico de uso múltiple. Tendrá un espejo primario de 2,4 metros, capaz de concentrar la radiación electromagnética en la totalidad de la parte óptica del espectro. Desde un comienzo estará equipado con un conjunto de instrumentos científicos que cumplirán las misiones siguientes: registrar imágenes astronómicas de resolución extremadamente alta, detectar objetos extremadamente débiles, recoger diversas clases de datos espectrográficos y hacer medidas muy precisas de la posición de fuentes radiantes en el cielo. Las observaciones se harán desde una altura de unos 500 kilómetros, muy por encima de las capas atmosféricas perturbadoras.

En el desarrollo de los objetivos del Telescopio Espacial han intervenido un nutrido grupo de científicos e ingenieros, que han venido trabajando durante casi un decenio bajo la supervisión de la Administración Nacional de Aeronáutica y del Espacio (NASA). Las empresas contratantes encargadas de la construcción son la Perkin-Elmer Corporation (responsable del propio telescopio) y la Lockheed Missiles and Space Company, Inc. (responsable de los sistemas auxiliares del vehículo espacial y de la integración de los componentes en un sistema que funcione correctamente). El coste total del proyecto se estima en la actualidad en 750 millones de dólares.

Se le ha asignado una vida útil de 15 años, aunque en principio no hay razón que impida su funcionamiento durante muchos decenios. Un elemento esencial para asegurar tan larga vida útil (y mantener los costos dentro de límites razonables) es la disponibilidad del transbordador espacial, que no se limitará a instalar el telescopio, sino que atenderá a su mantenimiento de una manera regular. Astronautas del transbordador visitarán el Telescopio Espacial siempre que los instrumentos a bordo del observatorio necesiten repaso, reparación o sustitución. A intervalos más dilatados (cada cinco años quizá) se traerá a la Tierra el Telescopio Espacial entero para restaurar el espejo y demás componentes delicados. El telescopio será después devuelto a su órbita.

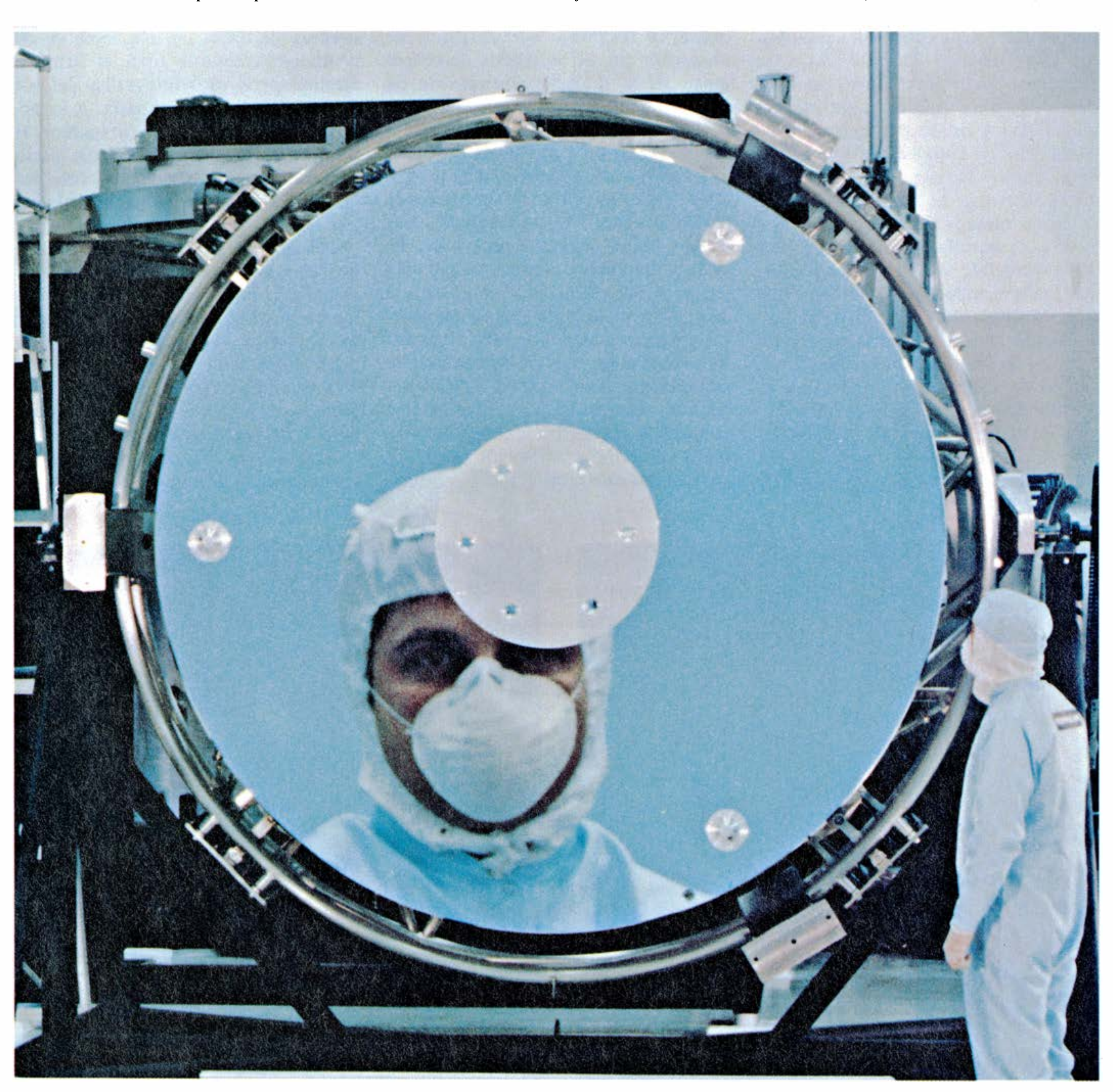
Dotado de la instrumentación adecuada, responderá a ondas electromagnéticas cuyas longitudes de onda varíen desde 115 nanometros (milmillonésimas de metro) en la región ultravioleta extrema del espectro hasta el millón aproximado de nanometros (un milímetro), en el infrarrojo lejano. Así, la banda espectral accesible al telescopio puede extenderse sobre un intervalo de longitudes de onda que varíen en un factor de 10.000. Potencia que contrasta con la de los telescopios instalados en tierra, que tienen una visión clara de colores cuya longitud de onda varíe desde 300 hasta 1000 nanometros aproximadamente, es decir, en un factor que no llega a 10.

Al hallarse inmune a los efectos degradadores de la turbulencia atmosférica, el telescopio podrá obtener imágenes de los objetos celestes mucho más nítidas que las de sus homólogos terrestres, incluso en las mismas longitudes de onda observables desde el suelo. La máxima resolución espacial que se alcanzará con el nuevo ingenio será de una décima de segundo de arco; compárese ello con este otro dato: la mayoría de las imágenes astronómicas hechas con telescopios terrestres tienen resoluciones no muy superiores a un segundo de arco. Esta resolución, diez veces mejor, consentirá observaciones más detalladas de los objetos extensos. Se espera también que los astrónomos lleguen a ver estrellas unas siete veces más alejadas del sistema solar de lo que alcanzan con los medios actualmente disponibles.

El programa de observaciones para el Telescopio Espacial será administrado en nombre de la NASA por el Con-

sorcio Universitario para la Investigación Astronómica (AURA), en el que se integran 17 universidades y que nació para sacarle el mayor rendimiento posible a distintas instalaciones de la National Science Foundation, entre ellas los Observatorios nacionales de Kitt Peak en Arizona y Cerro Tololo en Chile. El centro de la elaboración y análisis iniciales de los datos procedentes del telescopio será el Instituto de Ciencias del Telescopio Espacial, centro que está construyendo el AURA en el campus de la Universidad Johns Hopkins. El primer director del Instituto es Riccardo Giacconi, que dirigió los grupos científicos de los dos satélites de rayos X *Uhuru* y *Einstein*, de gran éxito. Del funcionamiento del Telescopio Espacial se ocuparán, en responsabilidad mancomunada, el Instituto y el Centro de Vuelos Espaciales Goddard de la NASA en Greenbelt, Maryland. El Centro Goddard ejercerá el control directo del satélite y servirá de punto colector de los datos transmitidos desde aquél hasta la Tierra.

La Agencia Espacial Europea (ESA) cubre aproximadamente el 15 por ciento del coste del Telescopio Espacial. Poseerá su propio centro de análisis de datos en la oficina central del Observatorio Meridional Europeo en Munich. La ESA suministra los paneles solares que proporcionan energía eléctrica al observatorio, una cámara de alta reso-



ESPEJO PRIMARIO del Telescopio Espacial, fotografiado en la factoría de Wilton, Connecticutt, propiedad de la Perkin-Elmer Corporation, inmediatamente después de que se hubiera recubierto su superficie anterior de una película reflectora de aluminio de ocho diezmilésimas de milímetro de espesor, seguida de una capa protectora de fluoruro de magnesio de 2,5 diezmilésimas de milímetro de espesor. El espejo, construido de vidrio de sílice fundida, tiene un coeficiente de dilatación térmica extremadamente bajo, un diámetro de 2,4 metros y pesa unos 785 kilogramos. Está formado por un núcleo celular ligero de 25 centímetros de espesor intercalado entre dos placas, de 2,5 centímetros de espesor cada una. En el curso de los 28 meses que duró el esmerilado y pulido necesario hasta dar a la superficie la forma apropiada, que es la de un hiperboloide cóncavo, se eliminó de la placa anterior un kilogramo de material. La persona enmascarada que se ve reflejada en el espejo está de pie junto al fotógrafo a unos 18 metros del espejo. Otro técnico, también enmascarado y provisto de un traje especial para preservar la limpieza de la superficie del espejo, está a la derecha. El orificio que ocupa el centro geométrico del espejo a través del cual pasará la luz del espejo secundario del telescopio se encuentra temporalmente cubierto por una placa metálica.

lución para objetos débiles para su sección instrumental y cierto número de científicos y técnicos que se incorporarán al personal del Instituto de Ciencias del Telescopio Espacial. En contrapartida, los astrónomos europeos tendrán derecho al 15 por ciento del tiempo de observación. También trabajarán con el telescopio astrónomos de otras partes del mundo. Será, pues, un verdadero observatorio internacional.

Las primeras observaciones astronómicas desde el espacio se remontan a finales del decenio de 1940; se realizaron con cohetes V-2 capturados a los alemanes. Algunos de estos primitivos cohetes de combustible líquido llegaron a los Estados Unidos al terminar la segunda guerra mundial; se emplearon para enviar instrumentos científicos muy por encima de la atmósfera durante varios minutos de observación. Más tarde se desarrollaron cohetes menores de combustible sólido destinados, de una manera específica, a la investigación; elevaban una carga útil de alrededor de 45 kilogramos hasta una altura máxima de 160 kilómetros, dando un tiempo de observación, por encima de la influencia perturbadora de la atmósfera, de varios minutos de duración. El subsiguiente desarrollo de equipos electrónicos ligeros de estado sólido posibilitó la construcción de instrumentos científicos de capacidad y complicación crecientes para tales misiones, sin tener que aumentar prohibitivamente la potencia necesaria para elevarlos.

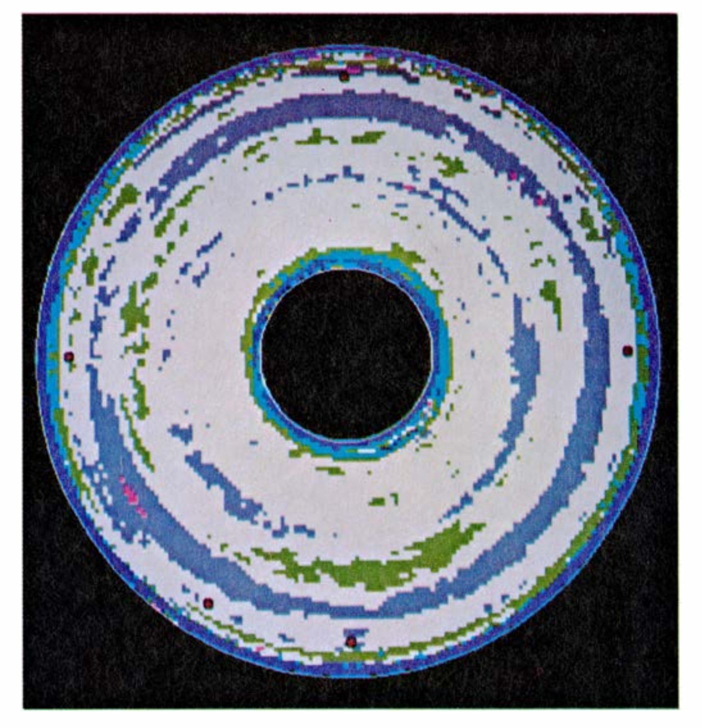
La primera aplicación de la tecnología de gran altitud fue el estudio del Sol. En 1946, un cohete elevaba un espectrómetro desarrollado por investigadores del Observatorio Naval de los Estados Unidos; se obtuvo el primer espectrograma solar en el ultravioleta, donde se revelaban características de absorción que no se habían detectado antes en la radiación procedente de ningún objeto celeste. Hasta 1957 no se pudo registrar la radiación ultravioleta de una estrella. La resolución espectrográfica de esta primitiva medida era tosca, con una anchura de banda de varias decenas de nanometros. Los cohetes de la primera generación no podían apuntarse con precisión; giraban libres en el espacio y ello impedía que suministraran la larga exposición necesaria para una medida precisa de la débil radiación procedente de una estrella distante. En el decenio de 1960 se elaboraron técnicas para apuntar los instrumentos transportados por cohetes hacia una estrella, utilizando pequeños giroscopios que proporcionaron un sistema de referencia inercial. Se registraron así espectrogramas estelares con una anchura de banda de una décima de nanometro, aproximadamente. Este éxito marcó el comienzo de una activa investigación acerca de muchos aspectos de las atmósferas estelares y de la materia interestelar.

Mientras tanto, otro grupo de astrónomos empleaba globos para elevar telescopios ópticos hasta unos 30 kilómetros de altura, por encima de las partes más densas de la atmósfera. A fines del decenio de 1950, un telescopio de 30 centímetros de este tipo, el llamado *Stratoscope I,* sacó fotografías del Sol de una nitidez extraordinaria. A lo largo de la década siguiente, su sucesor, el *Stratoscope II* de 90 centímetros, tomó varias fotografías de planetas y sistemas de estrellas con una resolución próxima a la décima de segundo de arco.

Los satélites artificiales, que podían permanecer en órbita años enteros, constituyen una plataforma para montar un telescopio óptico mucho mejor que un cohete suborbital o un globo. A medida que la tecnología aeroespacial ha progresado, los satélites se han convertido en los vehículos más idóneos para la astronomía extraterrestre.

Al igual que ocurrió con los primitivos cohetes y globos, las primeras ob-





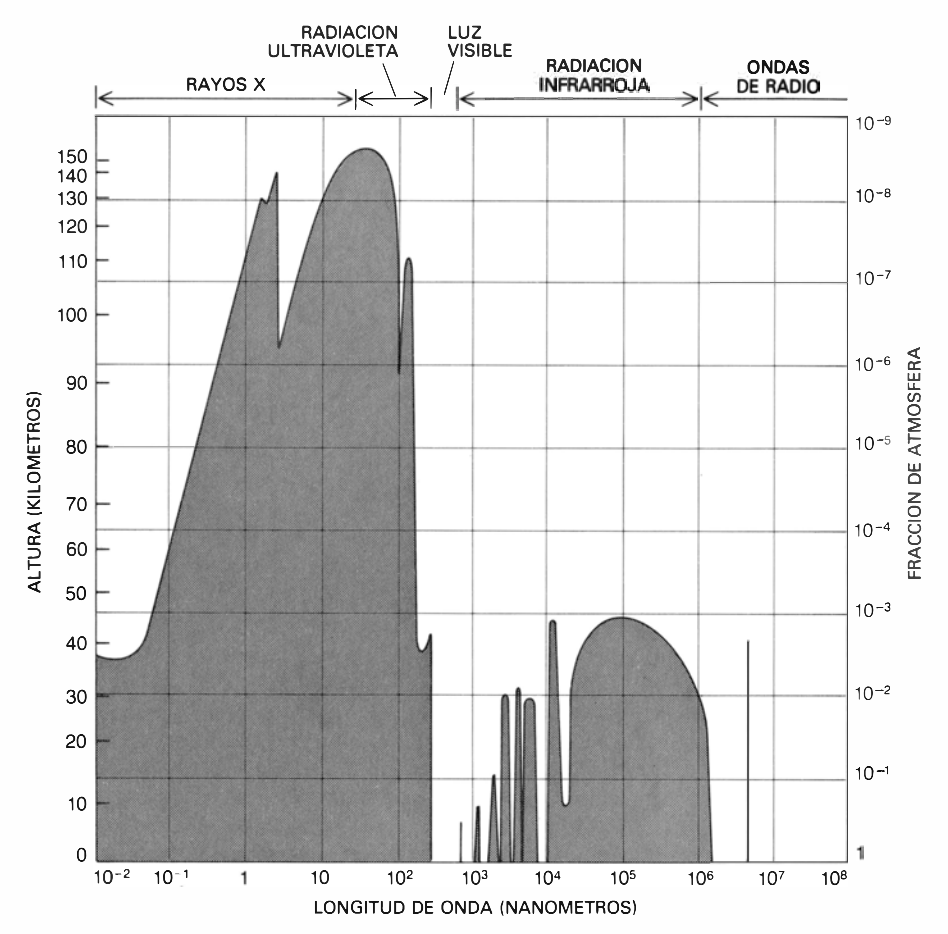
MAPAS TOPOGRAFICOS en falso color de la superficie del espejo primario, realizados en la pantalla de un terminal de ordenador con capacidad gráfica para ayudar a determinar la acción correctora necesaria en cada uno de los 24 ciclos del proceso final de pulido, que duró ocho meses. Los mapas se basaron en medidas interferométricas precisas de la forma de la superficie. Los dos mapas presentados se hicieron al principio y al final del proceso de pulido controlado por ordenador. Las zonas blancas representan la superficie media del espejo; las zonas en azul oscuro y rojo oscuro altas y bajas, respectivamente. Al principio, las desviaciones con respecto a la superficie prescrita llegaban a alcanzar 2,5 milésimas de milímetro (o micrometros); al final, la máxima desviación se había reducido a menos de la centésima parte de esa cifra en la mayor parte de la superficie del espejo. El espejo primario pulido es el mejor gran espejo astronómico jamás construido. Según Perkin-Elmer, "la superficie está tan cerca de la forma perfecta, que si el diámetro se aumentara hasta la anchura de los Estados Unidos continentales, no habría colina o valle que se separase de la superficie media en más de 60 milímetros".

servaciones que se hicieron con los satélites fueron del Sol. El proceso de hallar un objeto en el cielo y apuntar un telescopio hacia él es mucho más fácil con el Sol que con otra estrella menos cercana. Desde la década de 1960, la NASA ha construido y puesto en funcionamiento una serie de Observatorios Solares Orbitales, equipados con diversos instrumentos que permitieran estudiar su atmósfera.

Los primeros satélites NASA proyectados para observaciones estelares se designaron Observatorios Astronómicos Orbitales. Dos satélites de este tipo trabajaron satisfactoriamente, uno de 1968 a 1973 y el otro desde 1972 hasta 1981. Ambos se usaron, sobre todo, para analizar la radiación ultravioleta de las estrellas. El primero tenía una resolución espectrográfica bastante baja: su anchura de banda de medida era de 1,2 nanometros. El segundo, *Copernicus* de nombre, era muy superior en este punto: poseía una anchura de banda de 0,005 nanometros. El desarrollo de sistemas de guía precisos para tales satélites constituyó un notable éxito tecnológico. El telescopio del *Copernicus,* cuyo espejo medía 80 centímetros de diámetro, podía apuntar hacia una estrella durante varios minutos con una desviación máxima de 0,02 segundos de arco.

Los dos Observatorios Astronómicos Orbitales aportaron un verdadero aluvión de datos. Así, observaciones hechas con el *Copernicus* mostraron que gran parte del hidrógeno de las nubes interestelares estaba en forma de moléculas, no de átomos individuales. Se halló que muchos átomos de oxígeno de las regiones entre las nubes estaban altamente ionizados, lo que revelaba que el gas entre las nubes estaba muy caliente: del orden de un millón de grados Kelvin. Los datos de los satélites pusieron también de manifiesto que la relación cósmica de los átomos de deuterio, o hidrógeno pesado, a los de hidrógeno ordinario era de uno a 100.000, aproximadamente. De acuerdo con algunas teorías cosmológicas, esta medida apoya la idea de que el universo continuará expandiéndose por siempre.

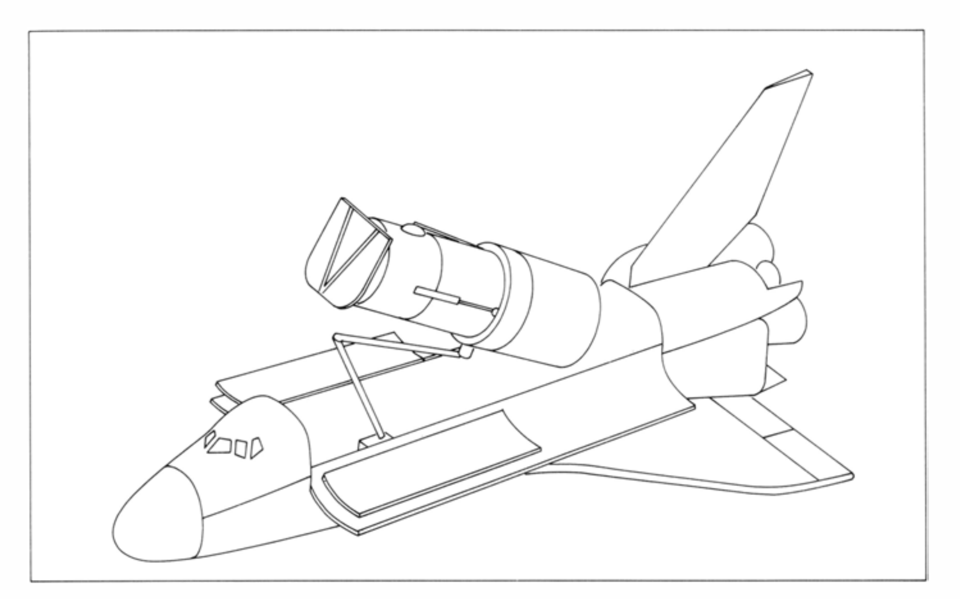
El último telescopio óptico en llegar al espacio ha sido el Explorador Ultravioleta Internacional, un satélite desarrollado conjuntamente por la NASA, la ESA y el Consejo Británico de Investigación Científica. Ha venido midiendo, desde 1978, el espectro ultravioleta de objetos bastante débiles. Aunque el rendimiento de este instrumento está limitado por el tamaño de su espejo (que tiene 46 centímetros de diámetro), se ha mostrado particularmente eficaz en la obtención de espectrogramas ultravioletas de núcleos de galaxias y en el análisis del gas interestelar en regiones remotas de nuestra galaxia.

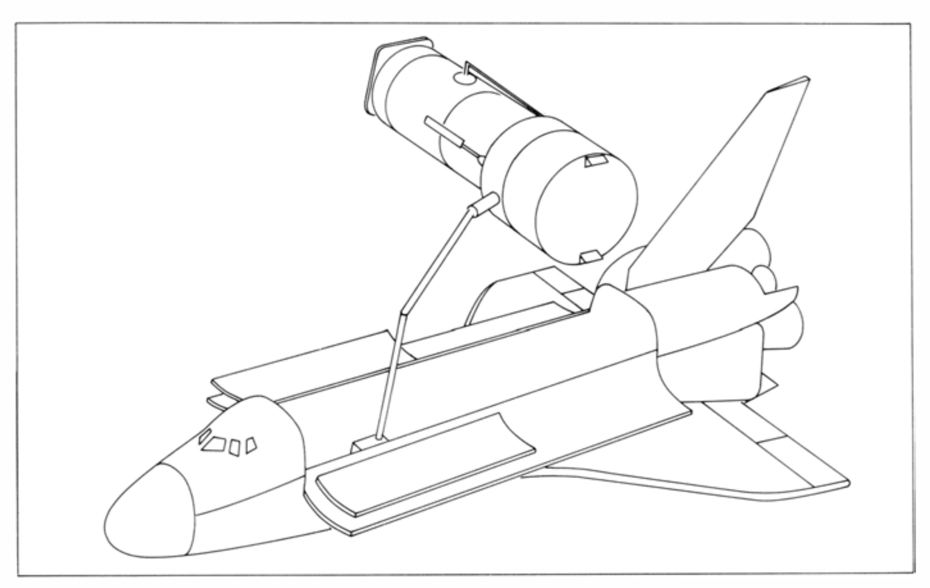


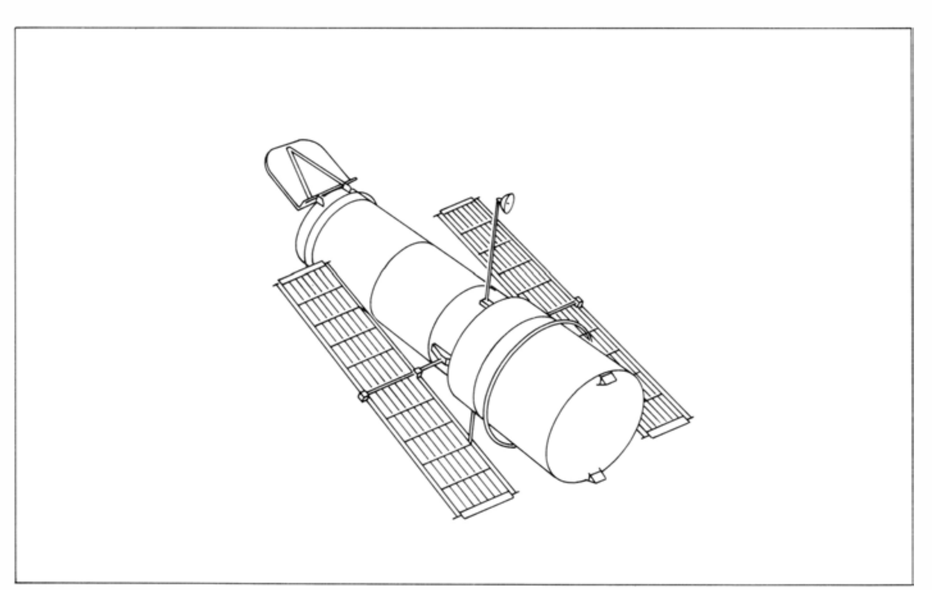
LA ABSORCION ATMOSFERICA de la radiación electromagnética limita la astronomía óptica desde la Tierra principalmente a la estrecha banda espectral correspondiente a la luz visible. La radiación en las vecinas regiones ultravioleta e infrarroja queda casi totalmente detenida. El borde superior de las zonas grises indica la frontera donde la intensidad de la radiación en cada longitud de onda se reduce a la mitad de su valor original. Un nanometro es una milmillonésima de metro, vale decir, 10 unidades Ångström.

La idea de un telescopio espacial mucho mayor fue tomando cuerpo lentamente a lo largo de los últimos veinte años. La primera noticia oficial de tal proyecto apareció en 1962 en el informe de un grupo de científicos, presentado a la NASA por la Academia Nacional de Ciencias, sobre el estudio del futuro de la ciencia espacial. El grupo recomendaba la construcción de un gran telescopio espacial como un corolario lógico del programa científico en ese campo de los Estados Unidos. La recomendación se repitió por un grupo de estudio semejante en 1965. Poco después, la Academia Nacional estableció un comité presidido por uno de los autores (Spitzer) para definir los objetivos científicos que cumpliría un telescopio espacial dotado de una abertura aproximada de tres metros. El informe de este último grupo se publicó en 1969. Pese a las muchas ventajas reseñadas para un telescopio espacial de semejante tamaño, la mayoría de los astrónomos estaban por aquel entonces demasiado embebidos en sus ocupaciones como para tomar parte activa en promover su desarrollo. La astronomía convencional (la de los telescopios instalados en la Tierra) acababa de entrar en una emocionante "edad de oro" con el descubrimiento de fenómenos del calibre de los quasars, la radiación cósmica de fondo en microondas y las estrellas de neutrones pulsantes; no había muchos científicos dispuestos a dedicar largos años de esfuerzo en el desarrollo de un equipo tan complicado y costoso como un gran telescopio espacial.

En 1972, otro comité de la Academia Nacional de Ciencias, presidido por Jesse L. Greenstein, del Instituto de Tecnología de California, programó el elenco de necesidades y prioridades de la astronomía para el decenio de 1970; de nuevo llamó la atención sobre las potencialidades de un gran telescopio espacial. Aunque la naturaleza y coste de tal instrumento apenas si quedaban por entonces esbozados, se consideró







EL TRANSBORDADOR ESPACIAL transportará el Telescopio Espacial hasta una altura aproximada de 500 kilómetros y lo pondrá entonces en órbita con ayuda del brazo mecánico. Los paneles solares de alimentación eléctrica, las antenas de comunicaciones y la puerta de apertura, todos ellos retraídos mientras el satélite es transportado en la bodega del transbordador, se desplegarán después de la puesta en órbita del satélite. El transbordador visitará periódicamente el telescopio para su mantenimiento, reparación y sustitución de piezas. Cada cinco años se traerá el satélite, de una pieza, para su restauración.

como un objetivo a largo plazo realista y deseable.

La NASA había reunido en el ínterin un pequeño grupo de astrónomos bajo la dirección de Nancy G. Roman con el fin de evaluar y discutir, desde el punto de vista científico, los estudios sobre la viabilidad del telescopio espacial que estaban realizándose en el Centro Goddard y en el Centro de Vuelos Espaciales George C. Marshall en Huntsville, Alabama. Aportaron su ayuda, en un comienzo, representantes de instituciones académicas, centros de investigación de la NASA e industrias contratantes.

En 1973, la NASA llamó a científicos de varias instituciones académicas para recabar su auxilio y opinión sobre el proyecto básico del telescopio y sus instrumentos. El grupo trabajó conjuntamente con los científicos e ingenieros de la NASA: se trataba de determinar qué objetivos eran viables y cúyo el orden de precedencia. El grueso de la orientación científica principal se debió a un grupo de 12 personas (en el que participaron los autores) presidido por C. R. O'Dell, de la Universidad de Chicago. Para dedicar toda su valía y fuerzas al proyecto, todavía sin financiación, O'Dell renunció a sus cargos de profesor y director del Departamento de Astronomía de la Universidad de Chicago y del Observatorio de Yerkes.

En 1977, la NASA convocó a un nuevo grupo de 60 científicos de 38 instituciones para participar en el proyecto y desarrollo del observatorio. Al frente de la dirección científica de este intento vuelve a estar O'Dell. En la composición actual del grupo de trabajo se aprecian personal de confianza de la NASA, jefes de investigación responsables de los instrumentos científicos iniciales, varios científicos interdisciplinarios (Bahcall es uno de ellos) y especialistas en tratamiento de datos, operaciones con vehículos espaciales y óptica de telescopios.

Faltó muy poco para que no se yugulara el programa del Telescopio Espacial. Entre 1974 y 1978, el proyecto estuvo en peligro constante de suspensión o aplazamiento indefinido a raíz de las correspondientes revisiones del presupuesto por parte del Congreso y la Presidencia. Tras un intenso período de contactos y visitas, en el que participaron no sólo cientos de astrónomos, sino también muchos científicos de otros campos, llegó por fin, en 1977, la autorización para su construcción. El programa pasó su primer examen presupuestario en el Congreso en 1978. Desde entonces, ha recibido en todas las

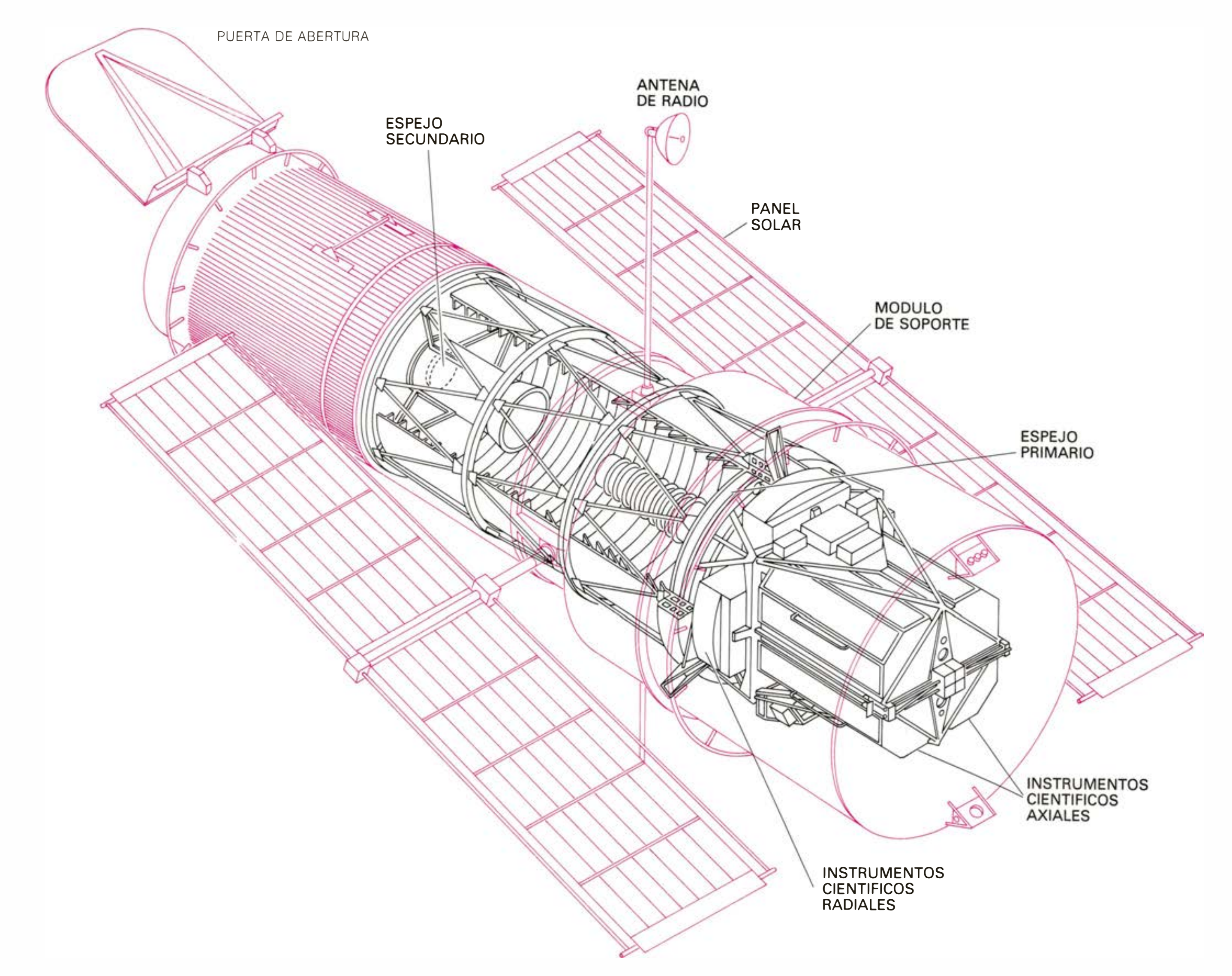
ocasiones una respuesta informada y favorable de la Colina del Capitolio.

Para cuando se recibió la aprobación formal, los estudios acometidos con todo pormenor por la NASA habían resultado en un proyecto completo, que en su mayor parte se está siguiendo en la práctica de la construcción del observatorio. Consta el telescopio de dos superficies reflectoras hiperboloidales: el espejo cóncavo primario de 239 centímetros y un espejo secundario convexo mucho menor montado a unos 5 metros delante del primario. La luz incidente en el espejo primario es reflejada hacia el secundario; éste la dirige a través de un orificio situado en el centro del primario; la imagen se forma en un foco que está a más de un metro por detrás del primario. Se trata, pues, de un modelo de telescopio perteneciente al tipo Ritchey-Chrétien de sistema óptico Cassegrain.

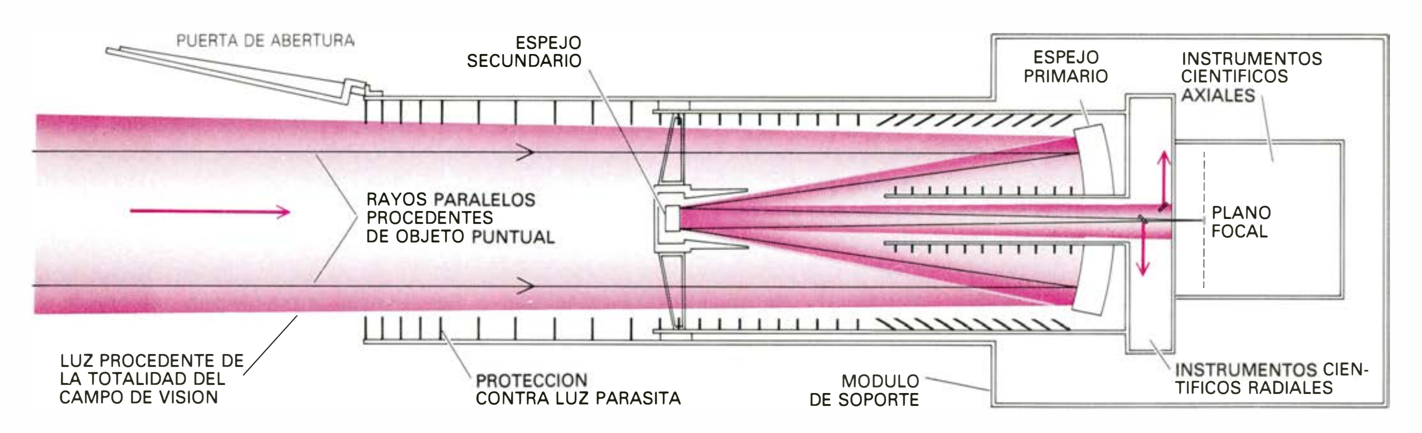
Los instrumentos científicos que detectan y miden la radiación concentrada en el plano focal se albergan en un conjunto de cajas montadas detrás del espejo primario. Hay cuatro cajas alineadas paralelamente al eje óptico del telescopio y otras cuatro dispuestas radialmente alrededor del eje. De las cuatro cajas radiales, tres acomodan el sistema de orientación fina del telescopio. El tubo de éste se extiende más de 3 metros delante del secundario a fin de proteger el sistema óptico de luz parásita, la mayor parte de la cual es luz directa del Sol y luz solar difundida por la Tierra y la Luna.

Un sistema de deflectores internos suministra protección adicional. El equipo electrónico y otros dispositivos están alojados en una sección toroidal que rodea al tubo del telescopio por su base. Hacia la parte media se despliegan dos paneles de células solares, que sirven para suministrar energía al equipo, y dos radio-antenas paraboloidales para comunicar con la Tierra. El cuerpo cilíndrico del satélite mide 14 metros de longitud por 4,5 de diámetro, en números redondos.

La característica más notable del Telescopio Espacial será la calidad sin precedentes de las imágenes formadas en su plano focal. Las superficies ópticas rayarán la perfección que puede alcanzar la moderna tecnología: la desviación media de las dos superficies reflectoras respecto a su forma ideal no excederá los 10 nanometros. Para evitar la distorsión térmica, los espejos están construidos de vidrio de sílice fundida, que tiene un coeficiente de dilatación térmica extremadamente bajo. Además, se mantendrán termostáticamente a temperatura casi constante mientras se hallen en el espacio. Por control remoto se ajustará la posición de los espejos, en mutua referencia y con respecto a la superficie focal, para



EN ESTA PERSPECTIVA DE CONJUNTO del Telescopio Espacial en su configuración desplegada, los componentes internos se han dibujado en negro y los externos aparecen en rojo. El cuerpo cilíndrico del ingenio tiene aproximadamente 13 metros de longitud y 4,25 de diámetro. Los instrumentos científicos se han construido de manera que puedan ser reemplazados por un astronauta provisto de traje espacial que trabajará desde el transbordador.



LA TRAYECTORIA DE LOS RAYOS LUMINOSOS en el Telescopio Espacial se dice que es "doblada": La luz procedente del espejo cóncavo primario se refleja en el espejo secundario convexo y atraviesa un agujero, en el centro del primario, antes de llegar al foco en el plano imagen, situado en la sección instrumental a varios decímetros por detrás del espejo primario. Técnicamente, el telescopio se llama tipo Ritchey-Chrétien del sistema óptico Cassegrain.

dar una imagen lo más definida posible. El sistema de orientación fina tomará una referencia en las imágenes estelares de la parte más externa del campo de visión del telescopio; se espera, pues, que logre mantener el eje óptico fijo dentro de 0,01 segundos de arco durante 10 horas al menos. (Unas ruedas de reacción interna servirán para apuntar el telescopio y mantenerlo fijo; al ordenar a una rueda de ésas que gire más deprisa en una dirección, el telescopio entero girará en la opuesta.)

Desde su lanzamiento y durante sus primeros años de funcionamiento, en la sección destinada a ellos habrá seis instrumentos principales. Son los cinco primeros: una cámara gran angular/planetaria, cámara para objetos débiles, espectrógrafo de objetos débiles, espectrógrafo de alta resolución y fotómetro de alta velocidad. Además, el sistema de orientación fina conferirá al telescopio capacidad astrométrica, esto es, capacidad para medir las posiciones exactas de las estrellas. Aunque los dos espejos tendrán un alto rendimiento de reflexión para la radiación a todas las longitudes de onda en la región óptica del espectro, no se incluirá ningún instrumento sensible a la radiación infrarroja en la fase preliminar. Sin embargo, todas las características del observatorio se han programado con la idea de acomodar una posible inclusión futura de un instrumento sensible a la radiación con longitudes de onda de hasta un milímetro.

Las aberturas de entrada de los cuatro instrumentos montados axialmente están en el plano focal del telescopio. Allí el campo total de visión, que mide 28 minutos de arco en unidades angulares, tiene casi medio metro de diámetro lineal; la escala resultante de la imagen en el plano focal es de 3,58 segundos de arco por milímetro. Mediante órdenes adecuadas de orientación, se puede dirigir la imagen de cualquier objeto en el campo de visión hacia cualquiera de los cuatro instrumentos axiales o hacia el quinto, montado radialmente. Cada instrumento está proyectado de modo que se puede retirar en órbita e instalar en su lugar un nuevo instrumento, siempre que lo haga un operador del transbordador espacial provisto del traje espacial adecuado.

Un ordenador de a bordo, independiente de los instrumentos científicos, controlará el funcionamiento del observatorio y tomará a su cargo el flujo de datos. El ordenador será reprogramable, posibilitando la reforma de los métodos a medida que se adquiera experiencia con los instrumentos. Los astrónomos y controladores de vehículos espaciales comunicarán con el Telescopio Espacial mediante el Sistema de Seguimiento y Transmisión de Datos de la NASA. Todos los datos se transmitirán a la Tierra mediante este sistema, para ser entregados, por último, al Instituto de Ciencias del Telescopio Espacial.

Los investigadores jefes de fila que habrían de hacerse cargo del desarrollo del conjunto inicial de instrumentos fueron elegidos tras una minuciosa y exigente selección. Para cuando se lance el satélite, cada investigador jefe, con sus respectivos colaboradores, habrán empleado más de ocho años construyendo un instrumento de aplicación general para uso potencial de todos los astrónomos. En recompensa a semejante esfuerzo, cada investigador principal y su equipo dispondrá de más de un mes de tiempo de observación.

El investigador responsable de la cámara gran-angular/planetaria es James A. Westphal, del Cal Tech. Este instrumento, como su nombre sugiere, puede operar de dos modos distintos: como cámara de gran campo o como cámara de alta resolución apropiada, entre otras cosas, para realizar observaciones planetarias. En cada modo, el sistema de detección está constituido por cuatro dispositivos de carga acoplada (CCD) o "pastillas" microelectrónicas de silicio que convierten una disposición de luz incidente en una secuencia de señales eléctricas. Cada "pastilla" es un cuadrado que mide casi un centímetro y cuarto de lado y está subdividida en una distribución de pixels, o elementos de imagen, con 800 de ellos por lado. Por tanto, cada pastilla alberga un total de 640.000 pixels; la imagen mosaico de cuatro partes formada por un conjunto de cuatro CCD tendrá más de 2,5 millones de pixels. Cada pixel suministra una señal eléctrica proporcional al número de fotones, o cuantos de radiación electromagnética, que inciden sobre él durante la exposición.

La cámara planetaria y de gran campo está montada en el lado del telescopio que generalmente quedará al otro lado del Sol. La luz incidente que pasa a lo largo del eje óptico del telescopio se dirige hacia el exterior en ángulo recto mediante un espejo plano desviador, sostenido por un brazo rígido y situado formando un ángulo de 45 grados con el eje óptico. El espejo diagonal desvía sólo la parte central del haz incidente; el resto de la luz rodea el espejo para llegar a los otros instrumentos.

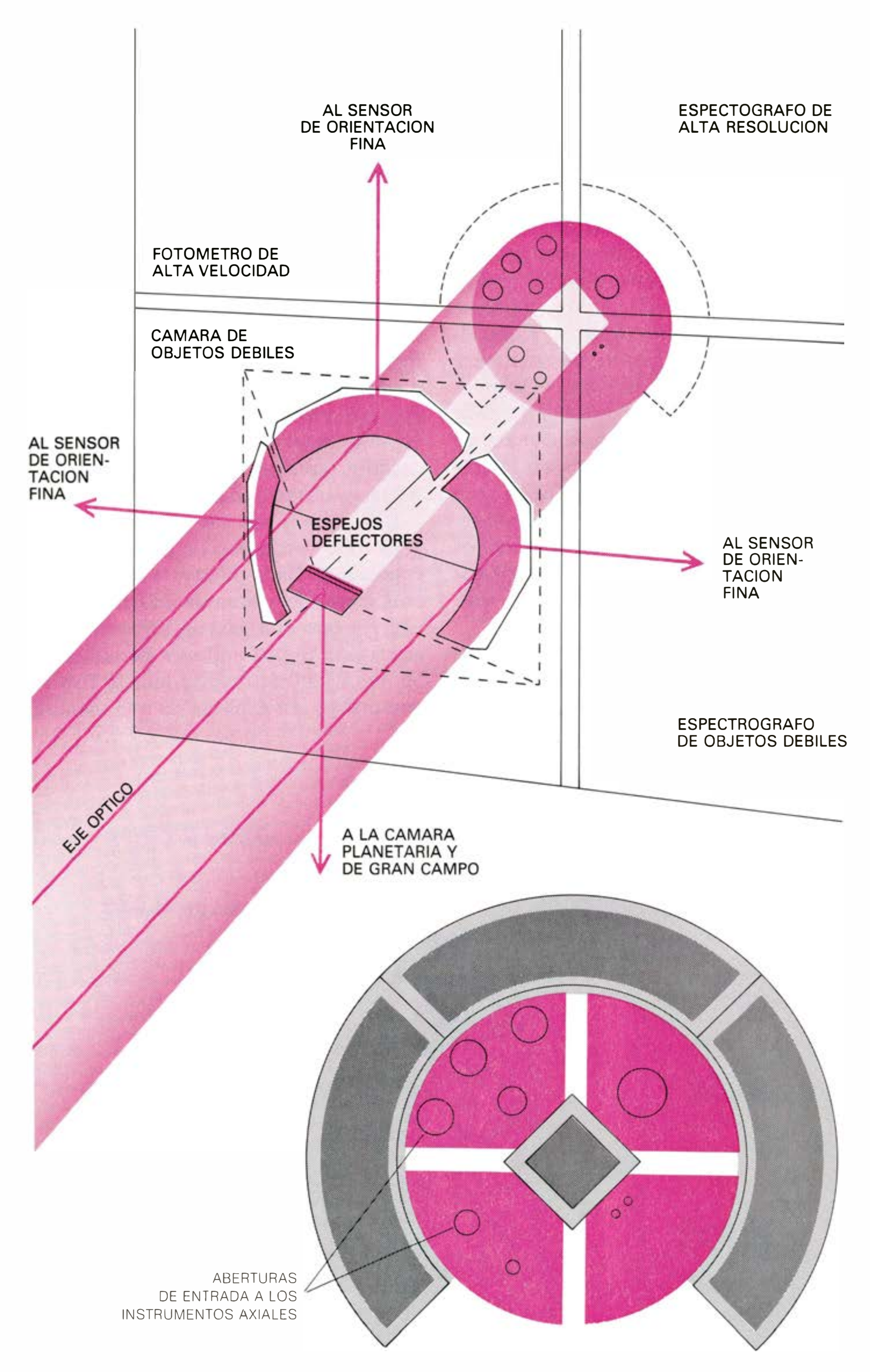
En el modo de gran campo, la cámara tiene un campo de visión cuadrado de 2,67 minutos de arco por lado, el mayor campo logrado hasta ahora en instrumento alguno. En este modo, cada pixel subtiende un ángulo de 0,1 segundos de arco. En cierto sentido, la cámara de gran campo compromete la resolución angular del telescopio a fin de proporcionar un campo de visión suficientemente grande que permita estudiar fuentes extensas; nebulosas planetarias, galaxias y cúmulos de galaxias.

Aun así, el campo de visión es mucho menor que el campo que se puede registrar en una placa fotográfica por un telescopio instalado en la Tierra. En el espacial, el campo queda limitado por el tamaño de los detectores microelectrónicos disponibles para la adquisición, almacenamiento y digitalización remota de imágenes. Los CCD para la cámara planetaria y de gran campo, suministrados por Texas Instruments, Inc., tienen más pixels que cualquier otro CCD usado en astronomía.

En el modo planetario, el campo de visión cuadrado de la cámara cubre, aproximadamente, una quinta parte del área del cielo que cubre en el modo de gran campo; en el modo planetario el campo mide 68,7 segundos de arco por lado; un pixel subtiende aquí un ángulo de 0,043 segundos de arco. La cámara planetaria aprovecha la casi totalidad de la resolución del sistema óptico, proporcionando al mismo tiempo un campo de visión que es más que adecuado para imágenes de disco completo de los planetas. La alta sensibilidad del sistema de detección CCD consiente el corto tiempo de exposición requerido para ciertas observaciones planetarias. Muchos astrónomos emplearán también el modo planetario para estudios de alta resolución de objetos galácticos y extragalácticos extensos.

La cámara planetaria y de gran campo sobresale entre los instrumentos del Telescopio Espacial por varios motivos. Recogerá, con gran diferencia, el mayor número de bits de información: más de 30 millones de bits por imagen. La respuesta espectral del detector será también la más amplia disponible entre todos los instrumentos del telescopio: la cámara será sensible a longitudes de onda que se extienden desde 115 nanometros, en la región ultravioleta lejana, hasta 1100 nanómetros en el infrarrojo próximo. La amplia cobertura espectral se ha conseguido recubriendo los CCD con un fósforo orgánico, llamado Coroneno, que convierte fotones de radiación ultravioleta en fotones visibles, detectables por los sensores de silicio. La excelente respuesta en el extremo rojo de la banda visible es atribuible a la natural sensibilidad de los CCD.

Los CCD empleados tanto en el modo de gran campo como en el modo planetario tienen un bajo nivel de "ruido" eléctrico de fondo; resultan, pues, idóneos para construir imágenes de fuentes débiles. Parte del ruido en un dispositivo de esta clase es térmico, y se reducirá enfriando termoeléctricamente los detectores hasta aproximadamente −95 grados Celsius. El calor genera-



LUZ INCIDENTE DESVIADA en distintas direcciones por un sistema de pequeños espejos deflectores situados cerca del centro de la sección de instrumentos del Telescopio Espacial detrás del espejo primario. El espejo plano en forma de diamante montado diagonalmente sobre el eje óptico dirige la luz hacia fuera hasta la cámara planetaria y de gran campo montada radialmente. Los tres espejos planos en forma de arco, dispuestos alrededor de la parte más externa del haz incidente, envían la luz a tres sensores del equipo de orientación fina, también montados radialmente. La luz que logra evitar esos cuatro espejos se reúne en el plano imagen en las aberturas de entrada próximas a la parte anterior de las cuatro cajas de instrumentos montadas axialmente. Las proyecciones de los espejos deflectores sobre el plano focal aparecen en gris oscuro, en la vista en plano de la parte inferior. Como los espejos deflectores interceptan el haz incidente a bastante distancia del plano focal, las zonas oscurecidas por éstos quedan ligeramente agrandadas; las zonas adicionales en penumbra están representadas por el enmarcado en gris claro de las zonas correspondientes a las proyecciones de los espejos. En el plano focal, el campo de visión es de 28 minutos de arco de diámetro angular. La cámara planetaria y de gran campo ve una región cuadrada de unos tres minutos de arco por lado en el centro del campo. El resto del campo hasta un radio de nueve minutos de arco se divide en cuadrantes, cada uno de los cuales es el campo de visión de uno de los cuatro instrumentos que están montados axialmente. La parte más externa del campo, entre los 9 y 14 minutos de arco del eje óptico, es explorada por el sistema de orientación fina, que tiene por misión no solamente apuntar el telescopio espacial sino también tomar mediciones precisas de la posición que están ocupando las estrellas.

do por el sistema de enfriamiento se disipará mediante un radiador que formará parte de la superficie exterior del satélite.

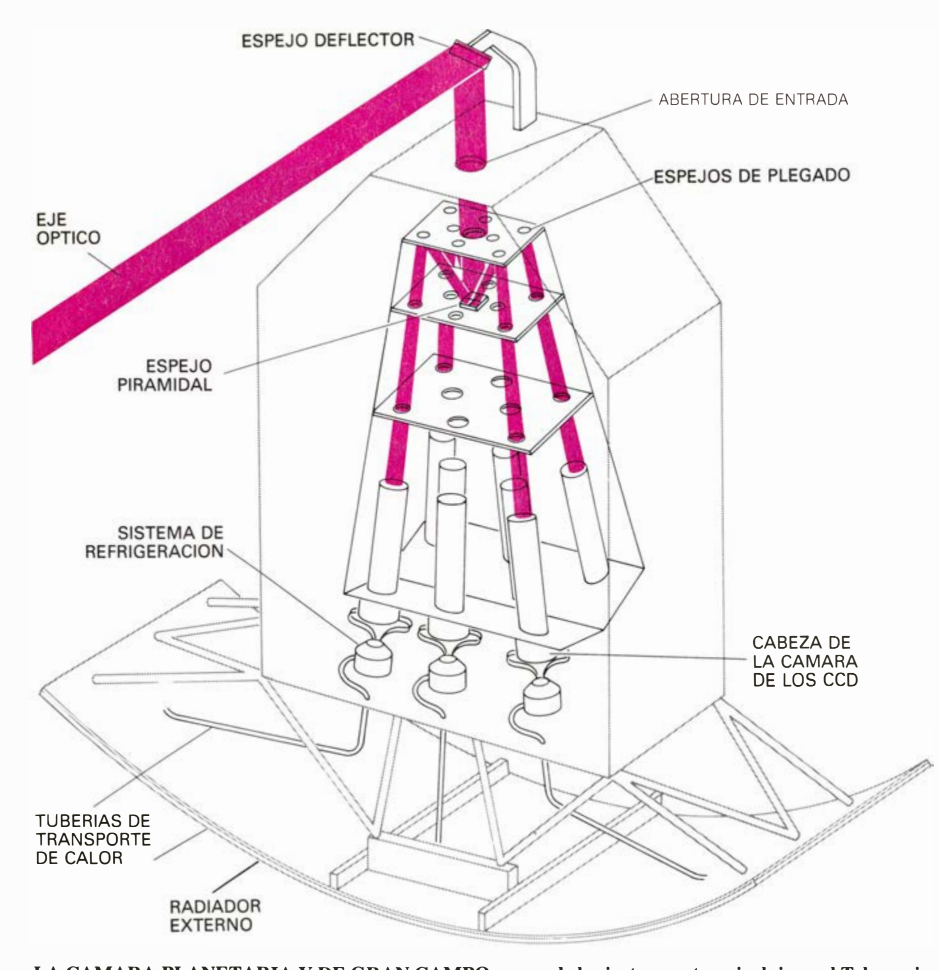
La luz incidente sobre el instrumento se puede dirigir o bien sobre los cuatro CCD de la cámara de gran campo, o bien sobre los cuatro CCD de la cámara planetaria, merced a un espejo piramidal que puede hacerse girar 45 grados alrededor de su eje, lo que permite alojar dos sistemas ópticos esencialmente independientes en un solo compartimiento de instrumentos. Cualquiera de los 48 filtros se puede intercalar en el camino óptico. Así, la cámara de gran campo y planetaria es un instrumento extremadamente versátil que sirve para una amplia variedad de necesidades astronómicas. Mencionaremos aquí sólo dos de las muchas investigaciones que se emprenderán con este instrumento.

En ambos modos la cámara se empleará para obtener una serie de imágenes de ciertas estrellas próximas con el propósito de averiguar si tienen compañeros planetarios. La decena de estrellas seleccionadas para el estudio lo han sido porque todas ellas tienen un movimiento propio grande (es decir, movimiento a través del cielo). Si alguna de ellas posee un sistema planetario, y merced a la extraordinaria resolución y precisa orientación del Telescopio Espacial, podrían detectarse "balanceos" periódicos en la trayectoria de la estrella producidos por la atracción gravitatoria de un compañero invisible. Las medidas son difíciles, pero el ingenio espacial está capacitado para resolver la vieja cuestión de si hay sistemas planetarios semejantes al sistema solar entre las estrellas próximas.

Los quasars son los objetos del universo más distantes y de mayor energía conocidos. Cada fuente compacta de ésas emite del orden de 100 veces más energía que una galaxia brillante integrada por 10.000 millones de estrellas. Se han propuesto varias teorías rivales que pretenden explicar de qué manera un quasar produce tan enorme cantidad de energía en un espacio tan pequeño; pero ocurre que determinadas comprobaciones observacionales, cruciales y necesarias para decidir sobre la cuestión, son irrealizables con instrumentos que tengan su base de asiento en el suelo. Parten algunas teorías de la idea de que los quasars constituyen galaxias "enfermas"; en otras palabras, se supone que los quasars representan un estado transitorio, valetudinario, en la evolución de una galaxia que, por lo demás, es normal. Para someter a prueba esas teorías se obtendrán imágenes de los quasars en alta resolución con la cámara planetaria y de gran campo para determinar si los brillantes objetos que aparecen como fuentes puntuales desde la Tierra están rodeados por la luz más débil y difusa de una galaxia. Hasta debería poderse decidir si el quasar representa una fase enfermiza de las galaxias jóvenes o de las antiguas. Esta pregunta fundamental no admite hoy contestación, debido a la falta de nitidez que presentan las imágenes obtenidas con los instrumentos de nuestras instalaciones terrestres.

La cámara para objetos débiles que proporcionará la ESA es uno de los cuatro instrumentos montados axialmente. Tiene por cometido principal aprovechar todo el potencial óptico del telescopio. Detectará con éste los objetos visibles más débiles y registrará imágenes que encierren la máxima resolución angular que se puede alcanzar con el sistema óptico. F. Macchetto, de la ESA, es el científico que dirige el proyecto de la cámara para objetos débiles.

Dicha cámara complementa en varios aspectos la cámara planetaria y de gran campo. La de objetos débiles tendrá una mayor resolución espacial, mientras que la segunda tendrá un mayor campo de visión. En la región espectral entre 120 y 400 nanometros, la cámara para objetos débiles conseguirá una imagen antes que la planetaria y de gran campo. Sin embargo, en las longitudes de onda mayores, hacia el rojo, la cámara planetaria y de gran campo será más rápida. Además de construir imágenes, la cámara para objetos débiles podrá determinar la polarización de la radiación detectada y tomar mediciones espectroscópicas de objetos puntuales y de objetos extensos. No se trata de duplicar el trabajo sino de asegurarse de que, siendo sus funciones muy parecidas, haya al menos en servicio una de ellas entre los instrumentos iniciales, en la hipótesis de que fallase la otra.



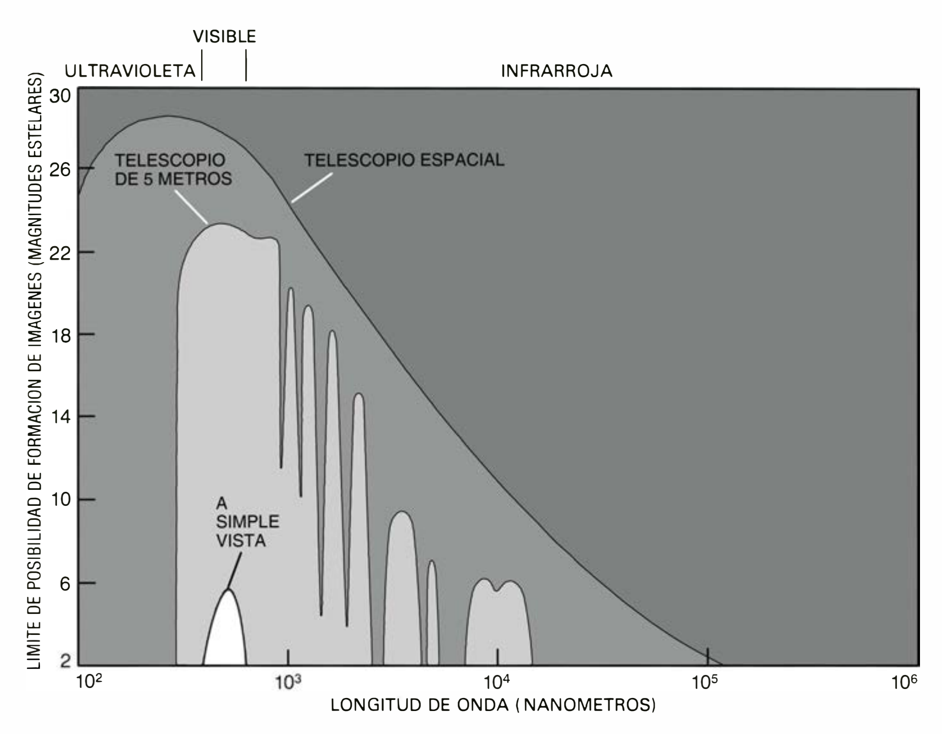
LA CAMARA PLANETARIA Y DE GRAN CAMPO es uno de los instrumentos a incluir en el Telescopio Espacial durante sus primeros años de funcionamiento. La cámara está proyectada para funcionar de dos modos. En cada uno de ellos, el sistema de detección consta de una distribución rectangular de cuatro "pastillas" de silicio sensibles a la luz, llamadas dispositivos de carga acoplada (CCD). La luz incidente, reflejada en el compartimiento de instrumentos montado radialmente por el espejo deflector diagonal, puede dirigirse hacia cualquiera de los cuatro CCD de la cámara de gran campo o los cuatro CCD de la cámara planetaria de alta resolución mediante un espejo piramidal que gira hasta 45 grados alrededor de su eje. En el camino del haz podrá introducirse un filtro cualquiera de entre 48 distintos. El radiador externo sirve para disipar el calor generado por el sistema de refrigeración asociado a los detectores.

En la cámara de objetos débiles se han previsto dos sistemas ópticos análogos e independientes para formar la imagen de una fuente puntual. Uno posee un campo de visión cuadrado muy pequeño, de 11 segundos de arco por lado; tiene un tamaño de pixel de sólo 0,022 segundos de arco. El campo de visión del otro sistema abarca 22 segundos de arco por lado y ofrece un tamaño de pixel de 0,044 segundos de arco. En ambos, el detector está constituido por un dispositivo intensificador de imagen semejante al tubo de rayos catódicos sensible a la luz en una cámara de televisión. A diferencia de los CCD de la cámara planetaria y de gran campo, un detector de esta clase cuenta fotones uno a uno.

La cámara de objetos débiles se ha proyectado de suerte que las imágenes de una fuente puntual producida por el telescopio sean exploradas por varios pixels. De aquí que sea el instrumento a elegir cuando se necesiten la máxima resolución posible y el máximo contraste contra el fondo del cielo. La cámara se hallará habilitada también para abordar estudios espectroscópicos y polarimétricos de objetos comparativamente débiles. Podrá enfocar objetos extremadamente débiles con un tamaño de pixel menor incluso (aproximadamente 0,007 segundos de arco).

Se espera que las tareas científicas de la cámara planetaria y de gran ángulo y de la de objetos débiles se superpongan. En razón de la resolución específica, campo de visión y región espectral requerida, un observador podrá elegir trabajar con una u otra. Mencionaremos aquí un tipo de observación en el que se recomienda el recurso a la cámara de objetos débiles.

Los cúmulos globulares son conjuntos esféricos de millones de estrellas que se ven desde la Tierra en una noche clara con un telescopio pequeño o incluso con gemelos. Como todas las estrellas del cúmulo se encuentran a la misma distancia del sistema solar, se pueden poner a prueba los modelos teóricos de la evolución estelar mediante el simple recuento del número de estrellas de cada clase en el cúmulo. La teoría clásica predice que cada cúmulo globular debe contener entre 10.000 y 100.000 de las estrellas llamadas enanas blancas. Estos objetos compactos representan la última fase en la evolución de las estrellas que han agotado su combustible nuclear, se han enfriado y han colapsado. En razón de su extrema debilidad, no pueden verse las enanas blancas a las grandes distancias de los cúmulos globulares con instrumentos instalados en el suelo. Pero la cámara de objetos débiles del Telescopio Espacial, sí detectará muchas de ellas en un cúmulo globular. Del estudio de sus propiedades sacaremos buena información para profundizar en el conocimiento de la evolución de las estrellas.



DOMINIO DE LONGITUDES DE ONDA potencialmente accesible al Telescopio Espacial. Se extiende desde la parte ultravioleta lejana del espectro *(izquierda)* hasta el infrarrojo lejano *(derecha)*. Para comparación, se indican también las bandas espectrales que pueden contemplarse a simple vista y con auxilio de un gran telescopio instalado en la Tierra (en este caso el telescopio Hale de 5 metros del Monte Palomar) bajo condiciones de observación normales. La escala vertical da el brillo relativo (en términos de la magnitud estelar) de los objetos celestes más débiles de los que se puede obtener imagen; un aumento de una unidad en la magnitud estelar corresponde a una disminución del brillo aparente de 2,5 veces.

El Telescopio Espacial dispondrá de dos espectrógrafos, o dispositivos ópticos que escinden la luz incidente de una fuente astronómica en haces separados de acuerdo con su longitud de onda. En espectroscopía, la resolución suele definirse como la relación de la longitud de onda de la luz incidente a la mínima separación que se puede medir entre dos longitudes de onda. Uno de los dos espectrógrafos a bordo del observatorio, el de objetos débiles, contemplará objetos estelares débiles con una resolución espectrográfica de 1000 (equivalente a medir una anchura de banda de una milésima parte de la longitud de onda). El investigador jefe del proyecto de este instrumento es Richard J. Harms, de la Universidad de California en San Diego.

El espectrógrafo de objetos débiles irá equipado con dos sistemas de detectores. Los dos son dispositivos llamados Digicon; uno es sensible a la luz roja y el otro a la azul y la ultravioleta. El sensor Digicon se basa en el efecto fotoeléctrico. La luz incidente es dispersada según su longitud de onda mediante una red de difracción, e incide sobre la superficie de una delgada capa fotocatódica depositada sobre una placa transparente. La luz de determinada longitud de onda alcanza una posición particular a lo largo del fotocátodo, produciendo la salida de un chorro de electrones conocidos por fotoelectrones. Un campo magnético focaliza los fotoelectrones hacia un punto cuya posición dependerá de aquella de donde emerja el haz en el fotocátodo, y por tanto, de la longitud de onda de la luz incidente. Los fotoelectrones se recogen en una distribución lineal de 512 diodos; cada uno registra la intensidad de la luz incidente a una longitud de onda particular.

El espectrógrafo de objetos débiles será sensible a la radiación de longitud de onda comprendida entre 115 y 800 nanometros. El instrumento tendrá, además, dos particularidades: será capaz de medir la polarización de la luz incidente y de detectar variaciones extremadamente rápidas (tal vez de algunos milisegundos) en el espectro de radiación emitido por fuentes brillantes. Habida cuenta de que la investigación de muchos problemas astronómicos depende del análisis espectral de la radiación procedente de objetos extremadamente débiles, se espera que el instrumento sea uno de los más activos en el Telescopio Espacial. Midiendo los espectros de quasars muy distantes, por

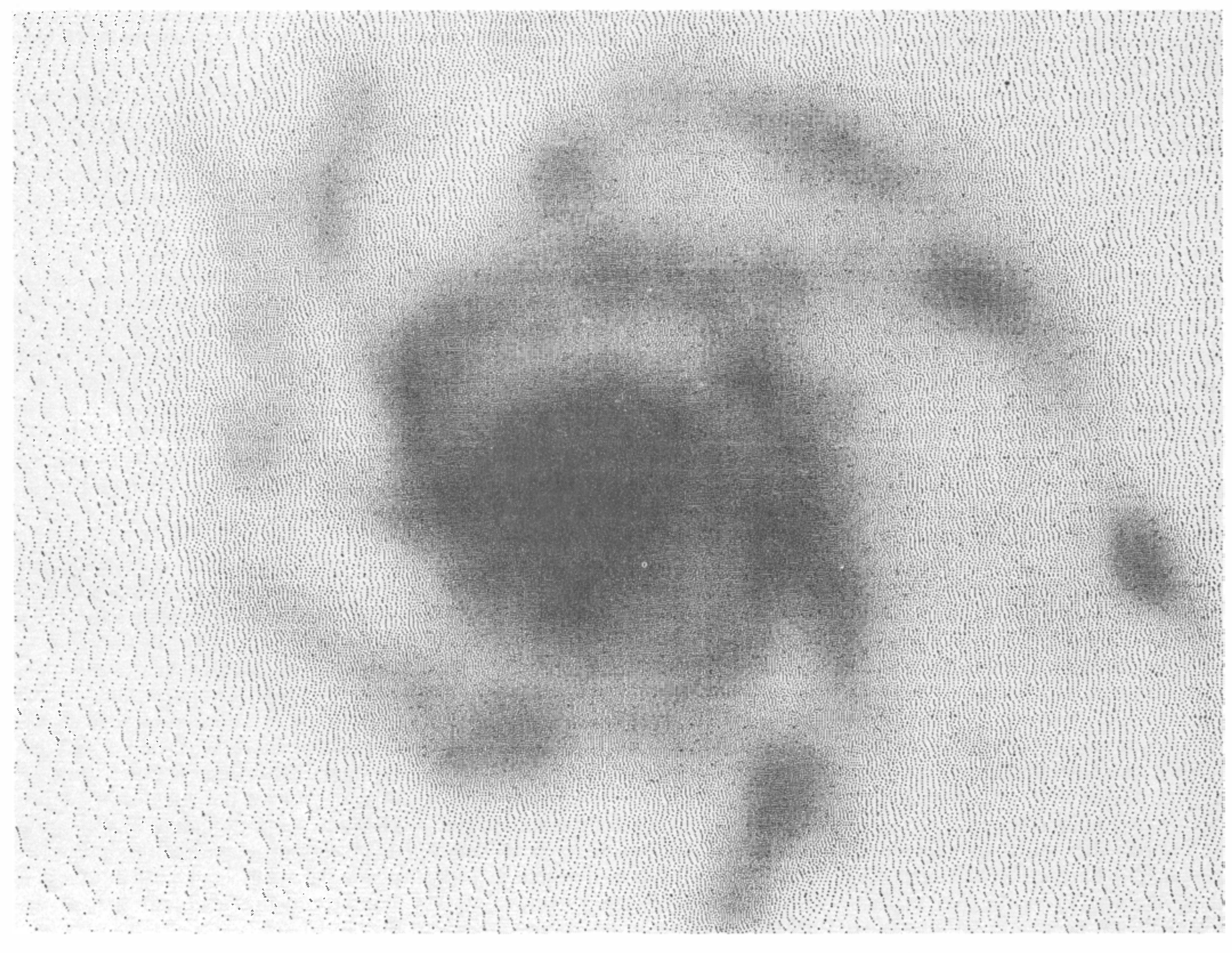
ejemplo, averiguaremos qué propiedades adornaban al universo de hace más de 10.000 millones de años, el 85 por ciento del tiempo transcurrido desde el origen (si, como supone el modelo cosmológico corriente de la gran explosión, el tiempo tuvo en realidad un origen). Se espera que los espectrogramas de los quasars más distantes indiquen la constitución de la materia en esa fase primitiva de la evolución del universo.

La investigación de ciertas cuestiones astronómicas requiere una resolución mayor que la que se puede alcanzar con el espectrógrafo de objetos débiles, porque la anchura de muchos rasgos de la emisión y la absorción es menor que la anchura de banda de medida del instrumento. El espectrógrafo de alta resolución satisfará esta necesidad. Bajo condiciones de funcionamiento normal, tendrá una resolución de 20.000. Las estrechas características espectrales que ni siquiera podrían detectarse con el espectrógrafo de objetos débiles, de resolución más baja, se medirán con precisión, informándonos de cuáles serían las condiciones físicas bajo las cuales se emitió la radiación. El espectrógrafo de alta resolución tendrá también un modo de operación en resolución ultra-alta, en el que la resolución espectrográfica se mejorará multiplicándola por aproximadamente 5 hasta alcanzar alrededor de 100.000. El investigador jefe a cuyo cargo corre el espectrógrafo de alta resolución es John C. Brandt, de Goddard.

Esa elevada resolución del segundo espectrógrafo se cobra su precio. Dividir el espectro en un número mucho mayor de bandas al objeto de calcular el flujo de fotones que discurre por cada banda comporta rebajar el número de fotones detectado en cada banda. Así, la mayor resolución se traduce en menor sensibilidad, y la mayor cantidad de información suministrada por el espectrógrafo de alta sensibilidad se puede obtener sólo para estrellas que brillen 60 veces más que las que pueden estudiarse con el espectrógrafo de objetos débiles. Esta diferencia corresponde a unas 4,5 magnitudes estelares. Para el modo de resolución ultra-alta, la diferencia de brillo corresponde a un factor de más de 300, equivalente a unas seis magnitudes estelares.

El espectrógrafo de alta resolución va equipado con seis redes de difracción intercambiables; cada red dispersa luz de diferentes longitudes de onda en direcciones distintas. Un espejo o red de la cámara forma entonces una imagen del espectro sobre la superficie emisora de electrones de un sensor Digicon. Girando el armazón circular sobre el que están montadas las redes, llevaremos cualquiera de ellas hasta el camino óptico del instrumento, lo que permite obtener lecturas espectroscópicas para cualquier longitud de onda entre 110 y 320 nanometros.

Este espectrógrafo, con su resolución normal, estará preparado para observar estrellas muy débiles, incluso las de magnitud 13, lo que viene a suponer



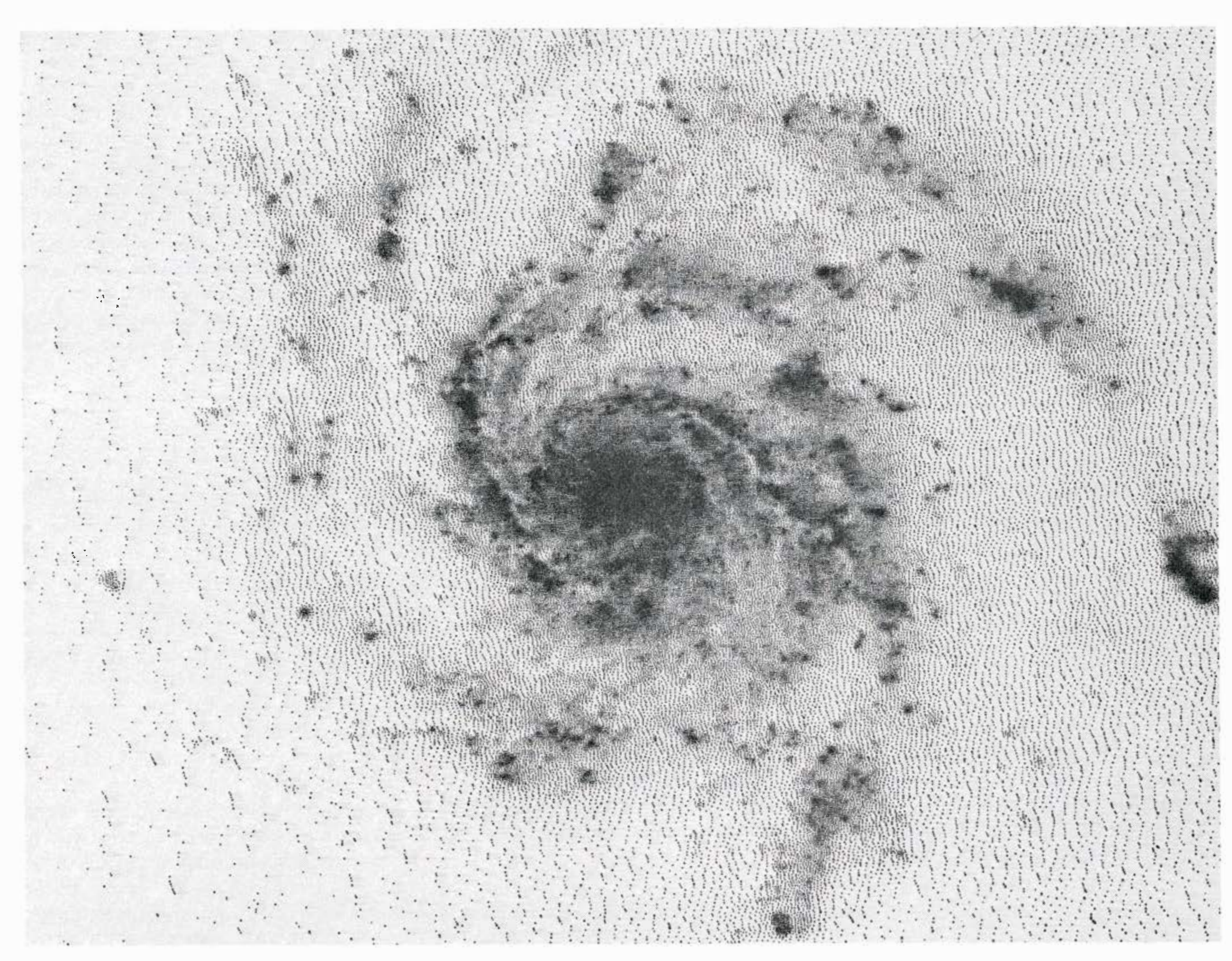
MEJORA DE DIEZ VECES en la resolución espacial: tal progreso se espera lograr con el Telescopio Espacial, lo que permitirá a los astrónomos realizar observaciones más detalladas de objetos extensos. En esta simulación, la fotografía de la izquierda representa la imagen de una galaxia espiral obtenida con el telescopio Hale de 5 metros y la fotografía a la derecha representa la correspondiente imagen obtenida con el Telescopio Espacial. En realidad, la

unas seis magnitudes estelares más débiles que las distinguibles con el telescopio *Copernicus*. Aunque no cabe hablar de un gran adelanto en lo que a sensibilidad se refiere sobre el espectrógrafo del Explorador Ultravioleta Internacional —unas cuatro magnitudes—, la resolución espectrográfica y la precisión fotométrica serán apreciablemente mejores en el instrumento del Telescopio Espacial.

La potencia de este instrumento abrirá interesantes líneas de investigación. El espectrógrafo de alta resolución facilitará, por ejemplo, el estudio del gas interestelar en puntos de nuestra galaxia y de otras donde ahora no se puede observar. Medidas preliminares realizadas por el Explorador Ultravioleta Internacional han mostrado que, en el gas en el "halo" galáctico entre la Tierra y la galaxia vecina más próxima (una de las dos nubes de Magallanes), hay átomos de carbono a los que faltan tres electrones, lo que indica que la temperatura en esta región es de alrededor de 100.000 grados Kelvin. Con el espectrógrafo de alta resolución se obtendrán datos mucho más precisos, que quizá revelen la relación entre este gas y la materia todavía más caliente detectada por el *Copernicus*. Las medidas que se tomen sobre cómo varían las propiedades de nuestra galaxia de un lugar a otro aportarán las claves que se necesitan sobre la evolución del sistema como un todo.

El espectrógrafo del alta resolución se aplicará también al estudio de nubes interestelares. Las observaciones de tales nubes hechas desde la Tierra apenas han podido detectar algunas rayas oscuras en el espectro, originadas cuando el gas de la nube absorbe la radiación procedente de las estrellas del fondo. En muchos casos, cada línea de absorción se divide en configuraciones múltiples, que pueden atribuirse a nubes distintas a lo largo de la misma visual. Las nubes se mueven con diferentes velocidades acercándose al sistema solar o alejándose de él, lo que altera las longitudes de onda características cuando absorben radiación. La separación de las líneas de absorción permite estudiar las nubes en sí mismas, siempre que sea suficiente la resolución del espectrógrafo. Con el de alta resolución llegaremos a analizar un amplio intervalo de características de absorción de varios átomos y moléculas y, por tanto, determinar las condiciones físicas de cada nube. Nuestro conocimiento sobre la manera en que tales nubes se agrupan y contraen para formar estrellas quizá dependa, ineludiblemente, de los resultados de tales estudios.

A Robert C. Bless y sus colaboradores, de la Universidad de Wisconsin en Madison, se les ha asignado el desarrollo del fotómetro de alta velocidad. Será su cometido realizar medidas de alta precisión, y con una resolución extraordinaria en tiempo, de la intensidad de la luz procedente de fuentes astronómicas a lo largo de un amplio intervalo de longitudes de onda. El fotó-



fotografía de la derecha es una versión digitalizada de una fotografía de una galaxia próxima tomada con el telescopio de 5 metros y la de la izquierda es una versión borrosa de la misma imagen obtenida desenfocando el original en una cuantía proporcional a la diferencia en la resolución efectiva que puede conseguirse con los dos instrumentos. Esta simulación fue preparada por John L. Tonry, adscrito al Instituto norteamericano de Estudios Avanzados.

metro podrá distinguir sucesos que disten en el tiempo 10 microsegundos escasos. A causa de las fluctuaciones de la atmósfera, resultan difíciles, si no imposibles, las observaciones terrestres de fuentes cuya variación en el tiempo sea de este orden.

El fotómetro de alta velocidad es el más sencillo de cuantos instrumentos conformen el grupo inicial que se instalará a bordo del Telescopio Espacial. Carece de partes móviles. Su éxito depende enteramente de la finura con que se pueda orientar el vehículo espacial para dirigir la luz de un blanco astronómico sobre una de sus 100 o más combinaciones de filtros espectrales y aberturas de entrada. El fotómetro posee cuatro detectores independientes, enfocados magnéticamente, que se llaman disectores de imagen; su funcionamiento recuerda al de los tubos fotomultiplicadores, salvo en que se puede limitar su respuesta a la presencia de fotoelectrones procedentes de la pequeña región del fotocátodo sobre el cual está incidiendo la luz. Cada disector de imagen se monta detrás de una placa que soporta un conjunto de filtros y aberturas de entrada.

La respuesta espectral global del disector de imagen se extiende de 115 a 650 nanometros. El instrumento está también equipado con un tubo fotomultiplicador sensible al rojo y un sistema para medir la polarización de la radiación ultravioleta con ayuda de uno de los disectores de imagen.

El fotómetro de alta velocidad podrá, en principio, detectar los menores objetos observables con cualquier instrumento del Telescopio Espacial. La posibilidad de llegar a distinguir entre sucesos separados en el tiempo por sólo 10 microsegundos implica, según la teoría especial de la relatividad, que se puedan detectar también variaciones en la intensidad de emisión de una estrella de diámetro no mayor de tres kilómetros: una dimensión lineal extraordinariamente pequeña para una estrella. De hecho, se acerca al diámetro que tendría el Sol si se comprimiese hasta la densidad que lo transformara en un agujero negro. Lo que nos lleva de la mano a otra peculiaridad: uno de los programas de observación previstos para el fotómetro de alta velocidad es la búsqueda de variaciones de suma rapidez en sistemas astronómicos de los que se sospecha alojen un agujero negro, en la esperanza de hallar ulteriores indicaciones de la existencia de esos entes escurridizos. El fotómetro de alta velocidad se usará también para observaciones menos exóticas; verbigracia: identificación de objetos ópticamente débiles observados sobre todo en longitudes de onda de radio o de rayos X.

Bajo las mejores condiciones de observación, las medidas de posición de cualquier estrella realizadas desde el suelo están limitadas por el tamaño del borroso "disco visible" de la estrella, que por lo general suele tener un segundo de arco de diámetro, al menos. En la determinación de la distancia angular entre dos estrellas es típica una incertidumbre de 0,1 segundos de arco, es decir, una décima parte del diámetro de la imagen estelar, para una sola observación. Promediando muchas observaciones, la incertidumbre se reduce a aproximadamente 0,01 segundos de arco. Los errores accidentales se traducen en las correspondientes incertidumbres en la determinación del paralaje de una estrella. (El paralaje es la variación angular media de la posición aparente de una estrella resultante de la revolución de la Tierra alrededor del Sol.) La determinación de distancias que trascienden el sistema solar se basa en gran medida en la determinación del paralaje de estrellas bastante próximas. Como la medida de una imagen estelar con el Telescopio Espacial se realizará con error inferior a 0,002 segundos de arco, la determinación de la posición de las estrellas, y por tanto del paralaje estelar, deberá ser cinco veces mejor que con los telescopios terrestres. Esa quíntuple mejora en la precisión de las medidas de paralaje estelar encierra una importancia fundamental en toda la astronomía estelar. Por ejemplo, conociendo la distancia precisa de ciertos cúmulos de estrellas bastante jóvenes de nuestra galaxia, los astrónomos determinarán el brillo absoluto de las estrellas en los cúmulos. A su vez, este conocimiento permitirá prolongar la escala calibrada de distancia, que se basa en la comparación de los brillos aparente y absoluto, hasta estrellas que se encuentran mucho más alejadas.

La razón de que se haya incorporado en el ingenio espacial un instrumento exclusivo para astrometría es la siguiente: el sistema de orientación fina, suficientemente preciso, realizará las necesarias medidas de la distancia angular entre estrellas. William H. Jefferys, de la Universidad de Texas en Austin, dirige el grupo de astrometría.

El tiempo de observación del nuevo y potente telescopio se repartirá entre astrónomos de todas las partes del mundo a través del Instituto de Ciencias del Telescopio Espacial, que tendrá la responsabilidad de facilitar el uso científico más eficaz. A fin de proporcionar a los astrónomos visitantes los sistemas de funcionamiento que rindan más, ayudar y aconsejar a los observadores sobre el mejor uso de los distintos instrumentos y crear una estimulante atmósfera de investigación, se está reclutando astrónomos de primera fila en los Estados Unidos y fuera de esa nación que deseen entrar en la plantilla del Instituto. Dedicarían la mitad de su tiempo a las necesidades del centro y la otra mitad quedaría disponible para sus propios programas de investigación. El nuevo instituto asesorará a la NASA acerca de las líneas generales de la política a seguir en asuntos relacionados con el Telescopio Espacial. Se pretende asegurar la participación de los astrónomos ajenos al Instituto en la política de éste a través de la creación de comités externos.

El Instituto solicitará propuestas externas para programas de observación específicos. Con ayuda de grupos de evaluación independientes, examinará las propuestas y seleccionará los proyectos más prometedores para su inclusión en el programa de trabajo del Telescopio. En muchos casos, los programas seleccionados se combinarán con los presentados por los grupos originales que se encargaron de los instrumentos científicos del telescopio, por otros miembros del grupo de trabajo del Telescopio y por los grupos europeos. El programa final y la preparación de la lista completa de instrucciones para el ordenador de mando será preparado por la NASA, que se reserva la responsabilidad del funcionamiento diario del observatorio.

Los astrónomos del Instituto aconsejarán a los astrónomos externos sobre la formulación de planes de observación. Otros astrónomos de la plantilla vigilarán el mantenimiento de la calibración de los instrumentos y de la elaboración inicial de los datos. Especialistas en ordenadores ayudarán a desarrollar los programas adecuados a usar por los astrónomos en el análisis de los datos. Finalmente, el Instituto de Ciencias del Telescopio Espacial pondrá en manos de los astrónomos los medios que necesiten a la hora de comunicar los resultados de sus estudios.

El Telescopio Espacial habrá de servir para resolver muchos rompecabezas astronómicos de carácter básico. Lo más emocionante, sin embargo, se producirá cuando las imágenes devueltas por el satélite revelen cosas que ninguno de la presente generación de astrónomos había soñado, fenómenos que quizá sólo la próxima generación tenga el privilegio de llegar a comprender.

Superconductores orgánicos

La capacidad de transporte sin resistencia de una corriente eléctrica había sido observada sólo en metales. La conductividad se aprecia ya en ciertos cristales orgánicos constituidos por moléculas planas que se hallan apiladas en zigzag

Klaus Bechgaard y Denis Jérome

La diversidad de compuestos orgánicos sintéticos que llenan la vida cotidiana surge de los recursos arquitectónicos inherentes al átomo de carbono. Los químicos, reuniendo cadenas o anillos de átomos de carbono, han sintetizado más de dos millones de substancias orgánicas que se emplean como fibras, recubrimientos, adhesivos, fármacos, gomas sintéticas, plásticos estructurales, etcétera. La experiencia colectiva en sistemas basados en el elemento carbono ofrece al diseñador de moléculas una extensa colección de materias primas y un arsenal de técnicas experimentales con los cuales se pueden crear nuevos materiales. Es como si se pudieran construir a medida moléculas orgánicas con casi cualquier propiedad física concebible.

Recientemente, nosotros y nuestros colegas hemos logrado crear una serie de materiales orgánicos con una propiedad rara y extraordinaria: la superconductividad, o sea, la capacidad de conducir sin resistencia una corriente eléctrica. Hasta ahora, la superconductividad había sido lograda sólo en materiales o aleaciones metálicas. La superconductividad constituye un fenómeno extraordinario incluso en esos materiales, la mayoría de los cuales son buenos conductores de la electricidad. La resistencia de un superconductor no es pequeña simplemente, sino exactamente cero; una corriente inducida en un superconductor no se disipa en forma de calor (como en los conductores corrientes), y por ello continúa circulando para siempre. La superconductividad en un sólido orgánico resulta un hecho más extraordinario todavía, por cuanto la gran mayoría de los materiales orgánicos sintéticos son aislantes de la electricidad.

Las condiciones bajo las cuales nosotros observamos por primera vez el estado de superconductividad en un compuesto orgánico fueron realmente extremas: a una temperatura de 0,9 grados Kelvin (o sea, –272 grados Celsius, aproximadamente) y presión de 12.000 atmósferas (12.000 veces mayor que la presión atmosférica en la superficie terrestre). Sin embargo, en menos de un año se encontraron cinco compuestos orgánicos sintéticos más que resultaron ser superconductores; uno de ellos es superconductor a la presión atmosférica, aunque la temperatura requerida continúe siendo baja para todos ellos. Existen buenas razones para creer que se podrá alcanzar la superconductividad a temperaturas superiores, tal vez incluso a 20 grados K. Y ahora que puede añadirse la superconductividad a la lista de propiedades observadas en las substancias orgánicas, se podrá asimismo aplicar todos los recursos de la síntesis química con el fin de moderar las condiciones bajo las cuales se mantiene el estado de superconductividad.

Desde su descubrimiento en 1911 por Heike Kamerlingh Onnes, de la Universidad de Leiden, la superconductividad se había considerado como un fenómeno de temperaturas bajas. Este descubrimiento fue consecuencia de un éxito tecnológico extraordinario: la licuación del helio. A la presión atmosférica, éste se condensa a 4,2 grados K. Kamerlingh Onnes empleó el nuevo líquido resultante como refrigerante, lo que le permitió estudiar la resistencia eléctrica de varios metales a muy bajas temperaturas. Normalmente, la resistencia de un metal decrece con la temperatura; y se había previsto, incluso a la luz de teorías premecanocuánticas, que si pudiera lograrse que un metal de estructura interatómica perfectamente regular pudiera alcanzar el cero absoluto de temperatura, entonces su resistencia eléctrica sería nula. Kamerlingh Onnes encontró, sin embargo, que la resistencia de un hilo de mercurio cesaba bruscamente a 4,2 grados K., temperatura muy baja, aunque por encima del cero absoluto. Poco después de su descubrimiento, los investigadores de Leiden observaron la superconductividad en otros metales: en el plomo, por debajo de 7,2 grados K, y en el estaño, de 3,7.

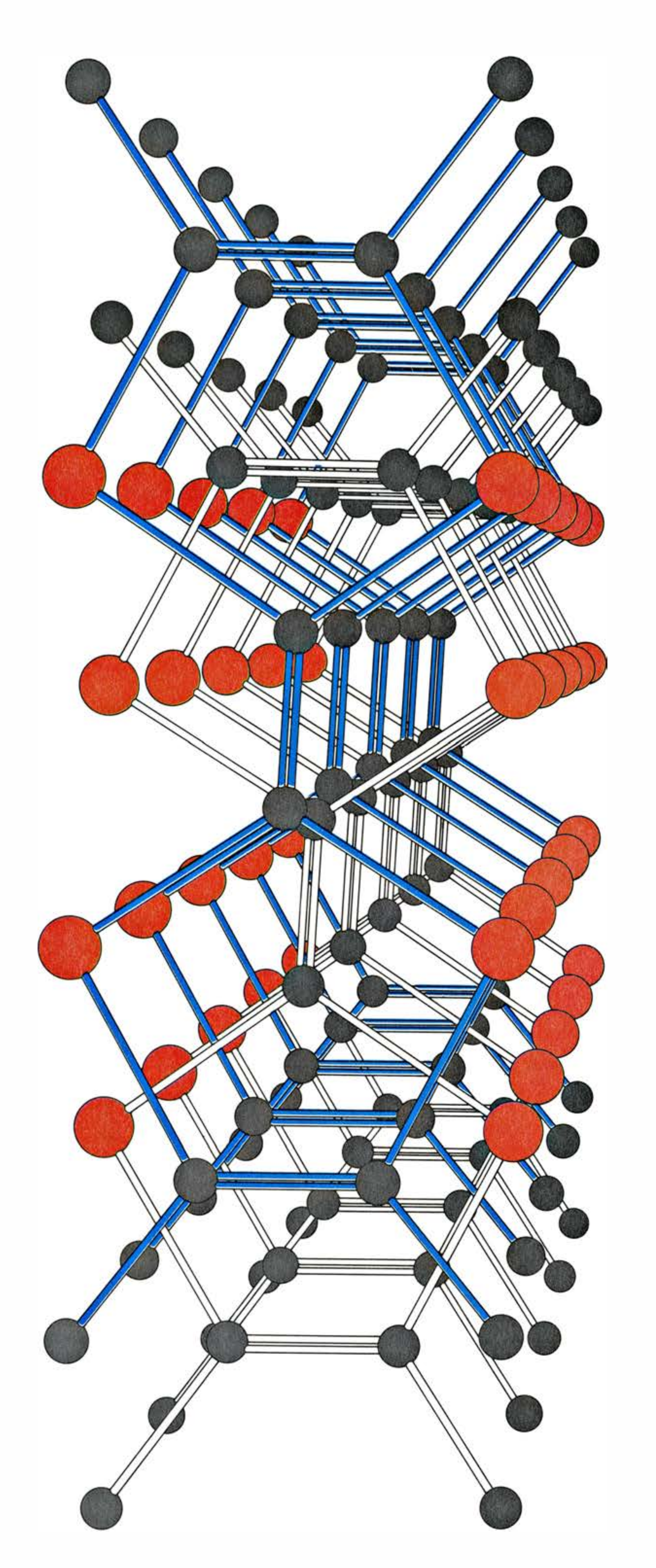
La temperatura en que aparece la superconductividad se denomina temperatura crítica. A medida que se descubrían nuevos superconductores, la temperatura crítica máxima medida aumentaba a razón de 0,3 grados por año. Se han observado temperaturas críticas relativamente elevadas en aleaciones metálicas, tales como de vanadio-silicio (17 grados), de niobio-estaño (18 grados) y de niobio-aluminio (18,7 grados). Así y todo, desde 1911, la temperatura crítica máxima tiende a estacionarse. El aumento más reciente de aquélla se registró en el año 1973 para una aleación de niobio y germanio (Nb_3Ge), resultando ser de 23,2 grados.

Desde el descubrimiento de la superconductividad, se hizo evidente que, si se pudiera lograr la superconductividad de materiales a temperaturas menos severas, su potencial tecnológico sería poco menos que ilimitado. Cabría entonces transmitir la energía sin pérdidas y lograr disponer de motores eléctricos más eficaces y electroimanes de potencial enorme. Se ha sugerido la construcción de supercomputadores que aprovecharan el pequeño calor liberado por elementos de circuitos superconductores. En efecto, la velocidad de un computador queda limitada, sobre todo, por el tamaño de éste, que determina la distancia media que una señal debe recorrer dentro del propio computador. A su vez, el tamaño de un computador está limitado por la capacidad de los elementos de sus circuitos para disipar el calor. Si la concentración de tales elementos es demasiado elevada, el calor que generan llegaría incluso a fundir el computador. Sin embargo, la aplicación de elementos superconductores

podría eludir tal problema. Por añadidura, tales elementos responderían a las señales con mucha mayor rapidez que los dispositivos semiconductores actuales.

La superconductividad se caracteriza también por un efecto magnético, llamado hoy efecto Meissner, observado por primera vez en 1933 por los investigadores alemanes K. W. Meissner y R. Ochsenfeld. Un superconductor colocado en un campo magnético genera su propio campo interno, que tiende a rechazar el campo exterior. Las líneas de fuerza del campo magnético externo se desvían entonces alrededor de la superficie del superconductor, en vez de atravesarla. (En ciertas clases de superconductores, las líneas de fuerza magnética pueden penetrar en la superficie del material hasta alguna profundidad.) El efecto Meissner sugiere asimismo una aplicación tecnológica: una pista construida con un material superconductor repelería el campo magnético generado por un vehículo, flotando así éste sobre un colchón magnético, y por ello se desplazaría prácticamente sin fricción alguna.

La búsqueda de la superconductividad a temperaturas altas ha sido por consiguiente una de las motivaciones en la física del estado sólido. Así y todo, y a pesar de los incentivos tecnológicos y económicos implicados, no se logró comprender el fenómeno de la



SE LLAMA PASILLO MOLECULAR a la pista preferida por los electrones de conducción en un cristal orgánico, uno de cuyos componentes moleculares es la molécula de tetrametiltetraselenafulvaleno, o TMTSF. Cada molécula de TMTSF posee una estructura plana que tiende a ceder sus electrones de valencia (los electrones más externos que se distribuyen por toda la molécula). Si se hace crecer el cristal de TMTSF con otra molécula o ion aceptor de electrones de valencia, sus moléculas se apilan en zigzag. En el diagrama se representa una vista desde arriba en el que alternan enlaces interatómicos azules y blancos en cada molécula. Puesto que las moléculas de TMTSF tienden a ceder electrones a las moléculas receptoras de los mismos (no se muestran en el diagrama), la pila de TMTSF posee una carga neta positiva. Los orbitales –o capas electrónicas– de cada molécula de TMTSF se distribuyen por encima y por debajo del plano molecular, de manera que los orbitales de moléculas adyacentes se solapan. De ahí que los electrones de conducción sólo puedan desplazarse fácilmente a lo largo del corredor formado por la pila, lo que explica que los cristales de TMTSF sean fuertemente conductores sólo en una dimensión. En el estado superconductor, los electrones de conducción forman sistemas denominados pares de Cooper, que se desplazan por el material en cuestión sin resistencia. Probablemente, los pares del TMTSF comienzan a formarse en los corredores moleculares.

superconductividad de los metales hasta casi medio siglo después de la primera observación de la propiedad. Ello requirió una aplicación profunda de los conceptos mecanocuánticos al movimiento de los átomos y los electrones en un sólido. Un acicate en esa búsqueda de la superconductividad a temperaturas altas fue la necesidad de progreso en el conocimiento de las propiedades de la materia en general. Y es de justicia reconocer que las investigaciones realizadas han contribuido muchísimo en lograrlo.

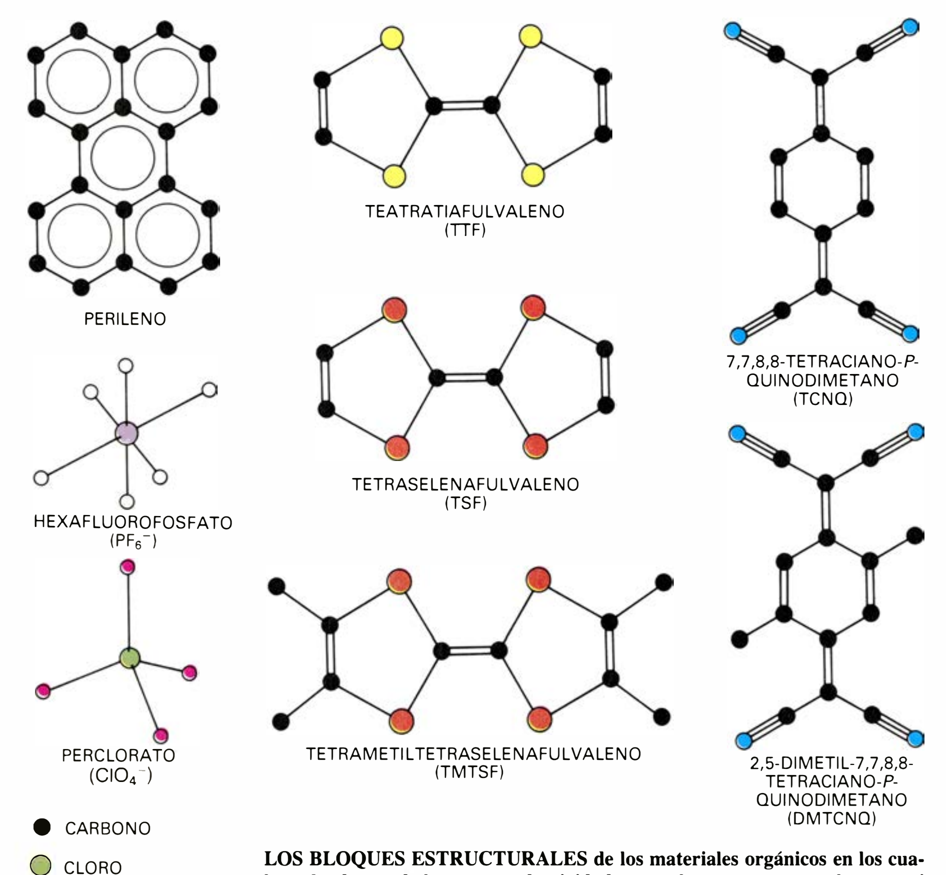
La teoría de la superconductividad en los metales fue formulada en 1957 por John Bardeen, Leon N. Cooper y J. Robert Schrieffer, a la sazón en la Universidad de Illinois, por cuya elaboración fueron galardonados en 1972 con el premio Nobel. De acuerdo con la teoría Bardeen-Cooper-Schrieffer, la superconductividad es un estado de la materia que surge cuando los electrones de un conductor se aparejan débilmente formando los llamados pares de Cooper. Los electrones de un par de Cooper se mueven con la misma velocidad y en sentidos opuestos. Aunque un par de Cooper está constituido por dos electrones, sus propiedades difieren de las que caracterizan al electrón lo suficiente para que sea tratado como una clase de partícula conductora totalmente nueva. La superconductividad es el resultado de un movimiento fuertemente coordinado de los pares de Cooper.

Puede parecer absurdo suponer que exista una atracción entre dos partículas con la misma carga eléctrica. Sin embargo, cuando los átomos se disponen en un sólido, los orbitales que definen la distribución espacial de los electrones más externos se solapan. Tales electrones, los llamados electrones de valencia, son los más débilmente unidos a sus átomos, tendiendo a vagar por los orbitales solapantes del sólido. Los átomos, que permanecen en posiciones fijas formando en retículo tridimensional, son de hecho iones cargados positivamente, puesto que han perdido por lo menos uno de sus electrones de valencia. La atracción entre los electrones de valencia apareados surge por mediación de tales iones positivos.

Cuando un electrón se desplaza por el retículo, las atracciones electrostáticas entre el electrón y los iones positivos genera una ondulación o deformación del propio retículo. El paso de un electrón cargado negativamente aproxima los iones positivos; las interacciones entre los iones crean unas fuerzas que tienden a restaurar la situación y por ello causan la vibración del retículo. Los movimientos vibratorios del retículo son mucho más lentos que el movimiento del electrón. De ahí que durante un tiempo relativamente largo, posterior al paso del electrón, los iones positivos permanezcan más juntos, creando así una región de carga positiva neta que atrae a un segundo electrón, formándose los pares de Cooper.

A medida que los iones vuelven a alejarse por acción de las fuerzas de repulsión mutua, se va creando una región donde la densidad de carga positiva es menor que la de su espacio circundante en el retículo, por cuyo motivo los dos electrones tienden a eludirla. A temperaturas altas ese efecto queda superado por las vibraciones térmicas aleatorias, pero puede surgir nuevamente por debajo de la temperatura crítica y así hacer que los electrones que forman cada par de Cooper actúen como si estuvieran unidos por un resorte. Los electrones constituyentes del par se ponen a vibrar de acuerdo con las variaciones de la densidad de la carga causadas por sus propios movimientos en el retículo atómico.

La conducción eléctrica en un superconductor resulta del movimiento neto de los centros de masa de los pares de Cooper, mientras que la conducción en un metal corriente es debida al movimiento neto de los electrones de valencia individualmente. Mas, para explicar el movimiento de los pares de Cooper en el retículo, es preciso recurrir a los principios de la mecánica cuántica. Cuando los electrones pasan al estado de unión, se minimiza la energía total del sistema si todos los pares de Cooper poseen el mismo momento. Si no fluye corriente alguna en el material, el momento de cada par de Cooper es cero, puesto que

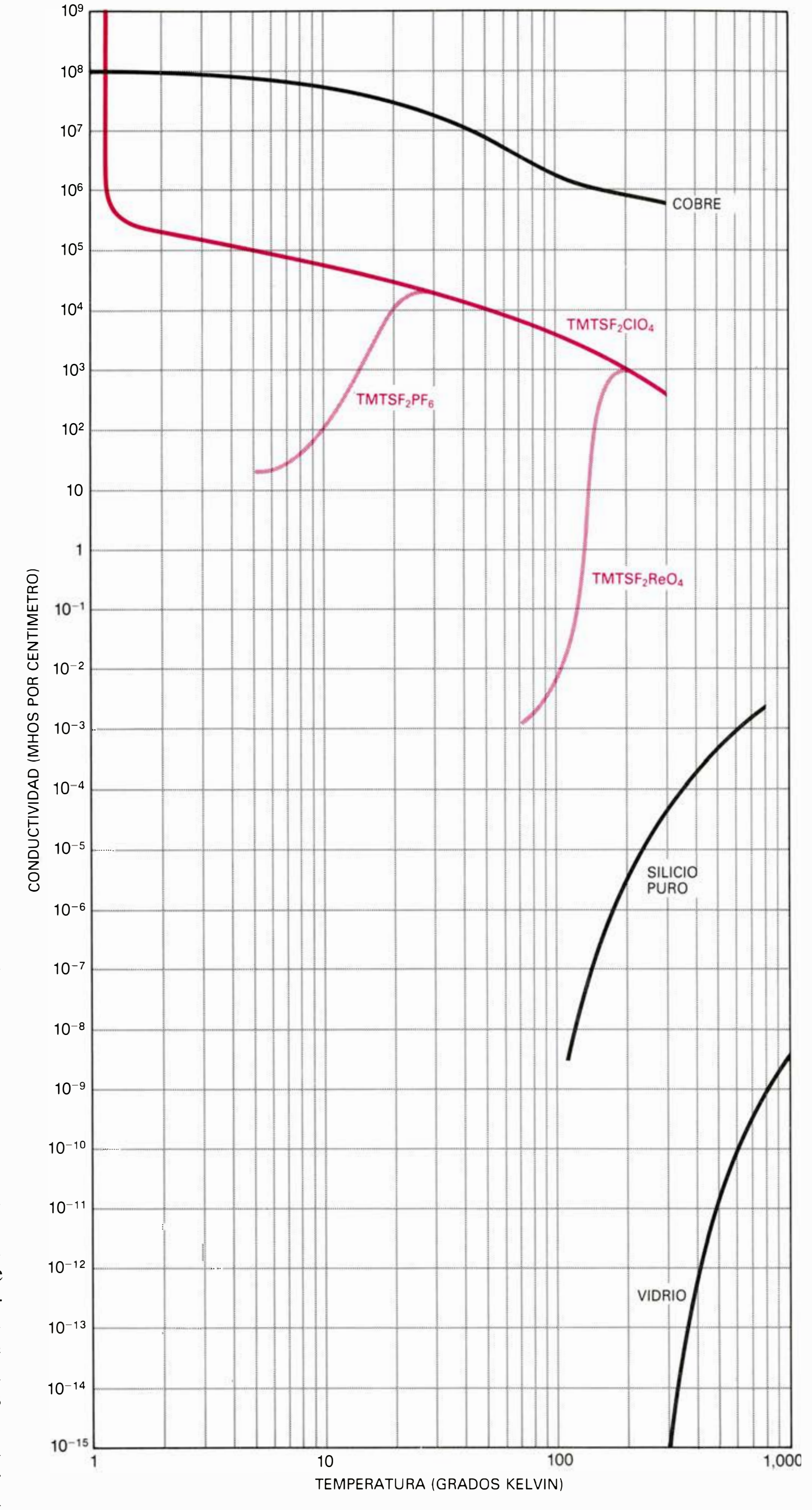


LOS BLOQUES ESTRUCTURALES de los materiales orgánicos en los cuales se ha detectado la superconductividad se acoplan en estructuras largas casi unidimensionales. Normalmente, en un solo cristal se combinan dos clases de moléculas. Puesto que las dos moléculas difieren en cuanto a sus afinidades por los electrones de valencia, una de ellas tiende a aceptar los electrones de valencia de la otra. La estructura resultante es una red de pilas de moléculas en la cual la carga alterna de una pila a otra. El bromuro de perileno fue la primera sustancia fuertemente conductora (no superconductora) que se descubrió. El TTS, el TSF y el TMTSF son moléculas afines, que ceden electrones; el TCNQ y el DMTCNQ aceptan electrones. Se han combinado los iones PF_6^- y ClO_4^- con el TMTSF, dando así los primeros superconductores orgánicos.

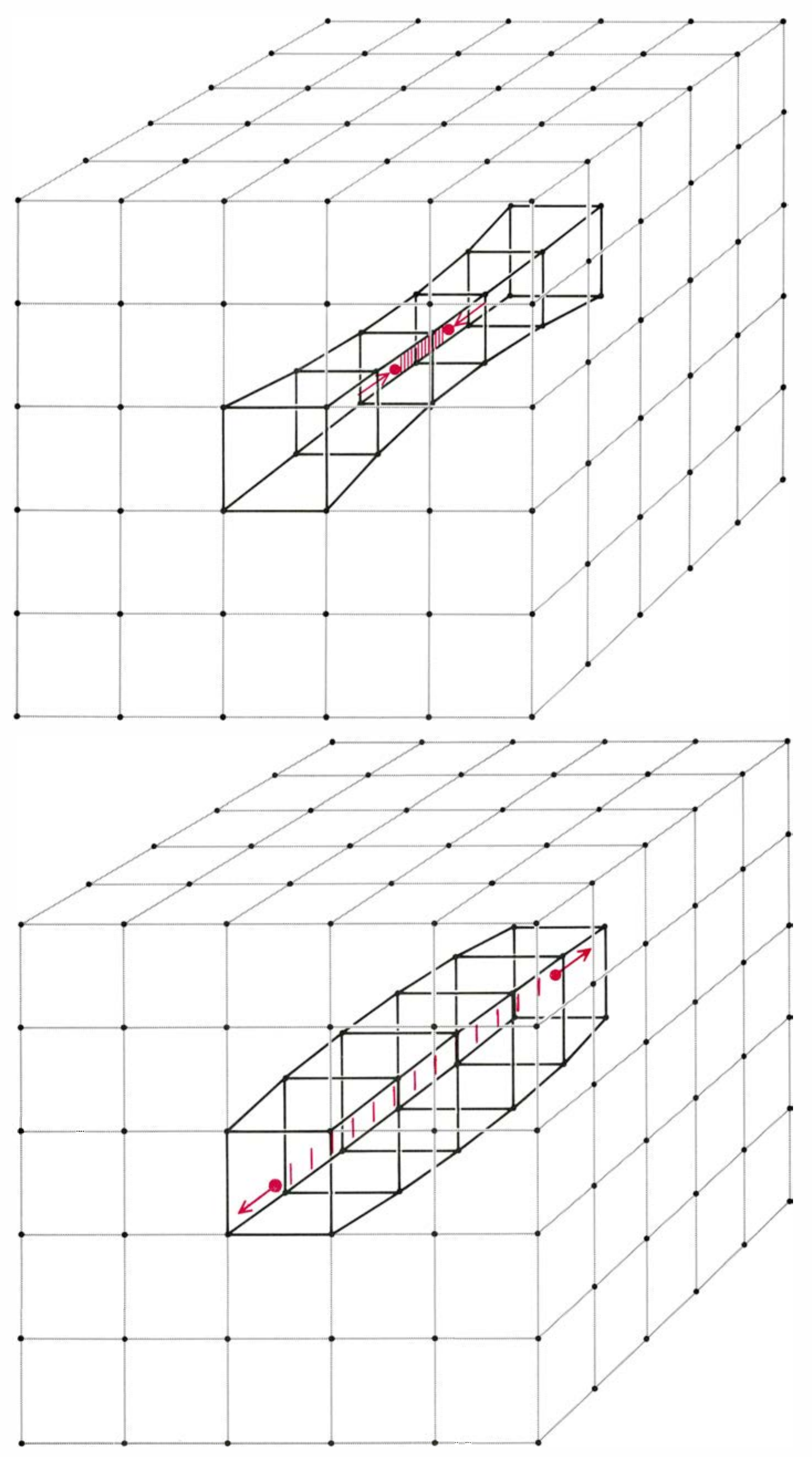
sus electrones se ven obligados a moverse en sentidos contrarios con respecto al retículo, y la velocidad del centro de masa de cada par debe ser cero exactamente. Cuando existe una corriente en el material, los electrones de cada par deben moverse de tal forma que los centros de masa de todos los pares de Cooper tengan un momento igual y constante. Como las manos de un funámbulo en su barra equilibradora, cada movimiento de un electrón que amenace romper el par de Cooper queda compensado automáticamente por su compañero, de suerte que siempre se mantenga el equilibrio.

¿Por qué un material que se opone al flujo de un electrón no se opone, en cambio, al flujo de los pares de Cooper? Normalmente, la resistencia eléctrica es el resultado de la dispersión de los electrones en movimiento por las imperfecciones del retículo atómico y por las vibraciones térmicas de éste a cualquier temperatura por encima del cero absoluto. La corriente o movimiento neto de los electrones a través del retículo resulta impedida por la dispersión; en consecuencia, la energía del sistema electrónico queda reducida en una proporción extremadamente pequeña. La dispersión de cada electrón y la disminución resultante de la corriente eléctrica se producen a causa de una pequeña variación de la energía vibracional almacenada en la red a una temperatura finita. La resistividad de un retículo metálico perfectamente regular sería teóricamente nula en el cero absoluto, porque a tal temperatura no existe en el retículo energía vibracional disponible para la dispersión de los electrones.

Por otro lado, puede anularse la resistividad electrónica de un superconductor a una temperatura definida, porque una vez que por la acción de un campo eléctrico los pares de Cooper se ponen en movimiento adquieren un momento neto. Este no puede perderse por dispersión. A diferencia de la dispersión de un electrón normal, que consume poca energía, la dispersión de los pares de Cooper requiere mucha más energía que la que mantiene los dos electrones juntos. Por añadidura, en un estado tan organizado como es el superconductor, un cambio en el momento de un par requiere asimismo un cambio de los momentos de todos los pares. De aquí que la energía necesaria para redistribuir los momentos de los pares de Cooper y establecer así la resistencia eléctrica es mucho mayor que



LA CONDUCTIVIDAD ELECTRICA es una de las magnitudes que varía más ampliamente. Los materiales pueden diferir en su conductividad por un factor superior a 10^{20}. (Si se incluye la superconductividad, entonces el factor es infinito.) Los conductores metálicos, como el cobre, aumentan su conductividad al disminuir la temperatura. Sin embargo, en los conductores orgánicos, los cambios estructurales o la aparición de propiedades magnéticas pueden convertirlos en semiconductores o en aislantes a bajas temperaturas. Se representa la variación de la conductividad de distintos materiales con la temperatura.



LA SUPERCONDUCTIVIDAD en un retículo metálico proviene de la interacción de los electrones de conducción con el mismo. Un electrón, cargado negativamente, desplazándose en un metal, atrae los iones cargados positivamente del retículo, y por ello éstos se aproximan entre sí a su paso. La región de carga neta positiva creada de esta forma tiende a atraer un segundo electrón, y, en consecuencia, el movimiento de éste resulta correlacionado con el del primero. Los dos electrones forman así un par de Cooper, actuando como si estuvieran conectados por un resorte. Se mueven a la misma velocidad, pero en sentidos opuestos. En las regiones con un exceso de carga positiva los dos electrones se acercan el uno al otro (*ilustración superior*). Las fuerzas originadas en el retículo, que tienden a restaurar la situación, actúan asimismo como resortes. Los iones positivos sobrepasan sus posiciones de equilibrio y empiezan a oscilar, creando efímeramente una región donde la densidad de carga positiva es menor que la del retículo circundante. Los dos electrones apareados eluden tal región y se apartan el uno del otro (*ilustración inferior*).

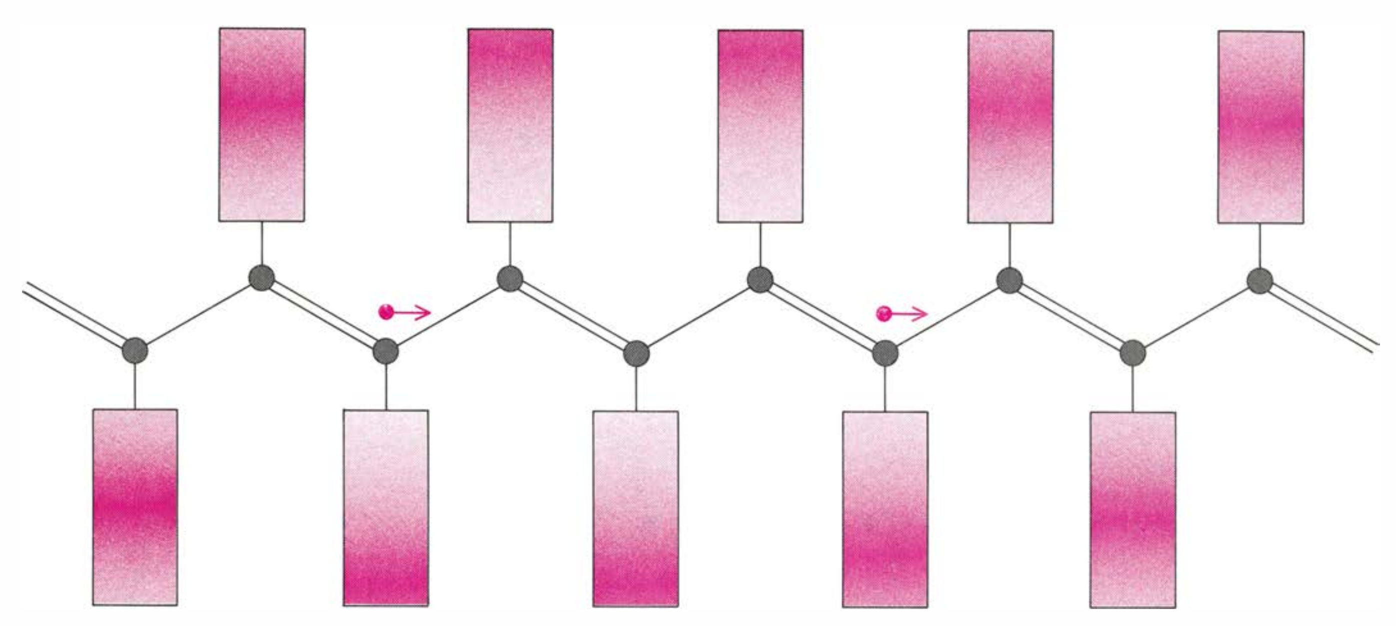
la energía vibracional disponible a baja temperatura en el retículo.

Por encima de la temperatura crítica no ocurre ya la unión de los electrones, porque entonces la agitación térmica de los propios electrones tiende a disociar los pares. Dicho de otro modo: si hubiera en el retículo suficiente energía disponible para redistribuir los momentos de los pares, bastaría esa energía para romperlos. Por debajo de la temperatura crítica, los pares continúan con su movimiento aparejado por el retículo sin cambio sensible durante años. Por encima de la temperatura crítica, los electrones sufren la dispersión como si de un conductor normal se tratara; la resistencia del material salta entonces de cero a un valor finito.

Puesto que la condensación de los electrones dando los pares de Cooper depende de las deformaciones del retículo atómico, la temperatura crítica de la superconductividad depende de la rigidez del retículo; es decir, de su resistencia a la deformación por el electrón en movimiento. Si los átomos del retículo tienen masas pequeñas y además los enlaces interatómicos son relativamente "blandos", las deformaciones son grandes en comparación con aquellas otras causadas por las vibraciones térmicas, y, por tanto, la temperatura crítica es elevada. Más exactamente, la temperatura crítica varía inversamente a la raíz cuadrada de la masa de los iones del retículo.

William L. McMillan, de los Laboratorios Bell, ha indicado no obstante que el ablandamiento del retículo puede establecer un límite superior para la temperatura crítica de la superconductividad. Cuando el retículo se ablanda, su frecuencia de vibración disminuye. Una menor frecuencia tiende a inhibir la formación de pares de Cooper, efecto que según sugiere McMillan pasa a ser dominante entre los 25 y los 30 grados K. Así pues, la teoría de McMillan puede explicar por qué sea imposible encontrar una substancia con una temperatura crítica superior a la del niobiogermanio.

En 1964, W. A. Little, de la Universidad de Stanford, propuso un mecanismo alternativo para la formación de los pares de Cooper, que no depende de las distorsiones reticulares [véase "Superconductivity at Room Temperature", por W. A. Little; SCIENTIFIC AMERICAN; febrero de 1965]. De acuerdo con él, podría fabricarse un superconductor con una cadena de moléculas orgánicas. Una cadena formada por átomos de carbono actuaría como una



PUEDEN FORMARSE los pares de Cooper en un superconductor orgánico por otro mecanismo diferente, basado en la polarización de las moléculas laterales unidas a la cadena de átomos de carbono. Cuando los electrones de conducción se desplazan por la cadena, repelen a los electrones de valencia de las moléculas laterales y tienden a concentrar su carga lejos de la cadena. En consecuencia, al paso de un electrón por la cadena se crea un exceso de carga positiva en la región más próxima a cada molécula lateral; por ello, un segundo electrón, que se desplace por la cadena, queda unido al primero. Una vez unidos, los dos electrones se ponen a oscilar como si estuvieran conectados por un resorte, igual que si de un metal se tratara. Las flechas señalan el movimiento de los electrones en el proceso de formación de la unión, antes de iniciar su oscilación. Este mecanismo ha sido propuesto por W. A. Little.

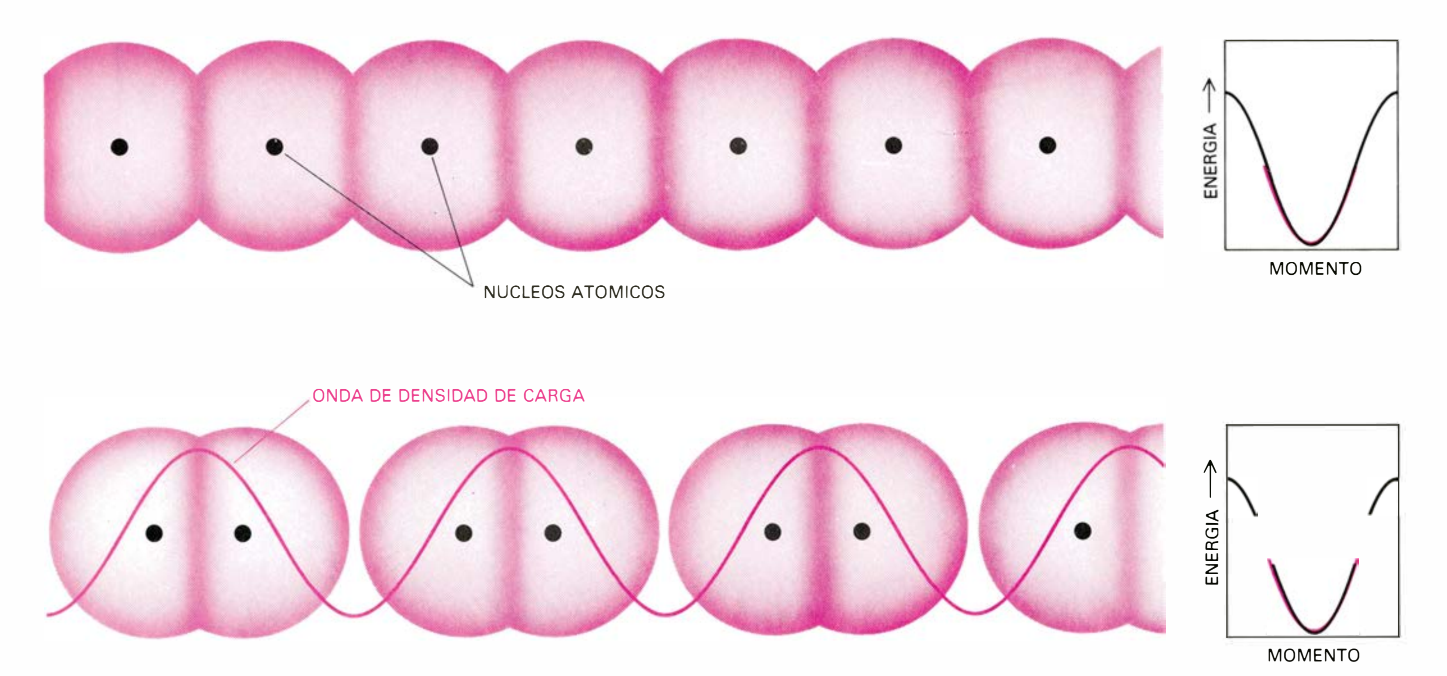
auténtica pista para los pares de Cooper, y las moléculas orgánicas unidas a tal pista representarían entonces el papel de retículo deformable. Little propuso que cada molécula orgánica fuera de hidrocarburo (concretamente, un colorante), con electrones de valencia débilmente ligados a ella. Tal molécula puede polarizarse sin dificultad, de forma que una parte de ella adquiere carga positiva mientras que la otra es negativa. Un electrón que se moviera a lo largo de tal cadena de átomos de carbono repelería los electrones exteriores de las moléculas hidrocarburadas, creando por tanto una región de carga neta positiva entre zonas de moléculas hidrocarburadas próximas a la cadena. Un segundo electrón que se moviera a lo largo de ésta sufriría la atracción de una carga eléctrica positiva relativamente importante, y por ello sería atraído indirectamente hacia el primer electrón.

Little ha argumentado que tal modelo sugiere la superconductividad a temperatura ambiente, indicando que, en principio, podría lograrse la superconductividad a temperaturas de hasta 2000 grados K. Según él, se podrían alcanzar temperaturas críticas tan elevadas porque la región de exaltación de la carga positiva al paso del primer electrón aparece en virtud del corrimiento de un electrón de la molécula hidrocarburada, en vez de un desplazamiento de un átomo en el retículo. Puesto que un electrón es unas 100.000 veces más ligero que un ión positivo de un metal corriente, la tendencia a formar pares de Cooper aumentaría aproximadamente en un factor igual a la raíz cuadrada de 100.000; es decir, de unos 300. Aunque el concepto de Little sigue siendo muy controvertido, ha despertado gran interés en cuanto al desarrollo de los superconductores orgánicos.

Nadie ha logrado construir aún una molécula cuyas propiedades se aproximen siquiera a las previstas por Little. Los primeros superconductores fueron, en cambio, el producto de unos intentos para construir, con materiales orgánicos, conductores eléctricos corrientes. El esfuerzo realizado está unido históricamente al desarrollo de componentes electrónicos fabricados con silicio y otros semiconductores. Los químicos orgánicos y los físicos del estado sólido buscaron armonizar el rango de propiedades eléctricas exhibidas por los elementos químicos para crear así, con sustancias orgánicas, lo que Little llamó una "tabla periódica fantasma" (*shadow periodic table*). Puesto que la mayoría de materiales orgánicos sintéticos, como el polivinilo o el polietileno, son aislantes eléctricos, se ha hecho muy difícil la búsqueda de conductores y semiconductores orgánicos.

La conductividad eléctrica es una de las propiedades que varían más ampliamente en la naturaleza: la de un buen conductor como el cobre supera, en más de 10^{20} veces, la de un buen aislante como el vidrio. Indicamos antes que la conducción normal depende del flujo neto, o desplazamiento, de los electrones de valencia a través del material. Bajo la influencia de un campo electromagnético, los movimientos esencialmente caóticos de los electrones de valencia quedan favorecidos ligeramente en una dirección determinada. En un átomo aislado, cada electrón de valencia ocupa un orbital discreto, que representa un estado con una energía determinada. No obstante, en una molécula o en un sólido, los principios de la mecánica cuántica no permiten que más de dos electrones posean exactamente la misma energía y el mismo momento. En consecuencia, cada nivel energético discreto de los electrones de valencia atómicos se descompone en muchos niveles ligeramente diferentes, distribuidos en una banda casi continua.

El desplazamiento de los electrones a través del material corresponde a un aumento de la energía electrónica total de éste, y sólo puede producirse fácilmente si los estados de energía permitida en la banda de energía de valencia no están totalmente ocupados. Los electrones de valencia en tal material pueden desplazarse por regiones del retículo en las cuales se solapan los orbitales atómicos parcialmente ocupados. Por otro lado, si la banda de valencia está ocupada completamente, la energía electrónica total puede aumentarse con sólo incrementar la energía de al-



PARA EXPLICAR la conducción unidimensional recurriremos a un sistema simple constituido por un conjunto de átomos de sodio dispuestos en formación lineal, cuyos orbitales de electrones de valencia se solapen. Cada átomo de sodio tiene un electrón de valencia; cuando los orbitales se solapan, los estados energéticos de los electrones de valencia forman una banda de niveles energéticos continua. Esta banda está semiocupada por tales electrones, y si se aplica un campo eléctrico exterior, algunos pasan a ocupar estados energéticos de la banda desocupados. El corrimiento de la distribución energética de los electrones corresponde a un movimiento neto de electrones a lo largo de los orbitales atómicos solapantes (*arriba*). Mas, por tratarse de un sistema unidimensional, su energía baja si los elementos del conjunto lineal no gozan de una distribución uniforme. La densidad de carga de los electrones de valencia en tal sistema puede variar con una periodicidad igual al doble del espaciado interatómico. Las distorsiones estructurales y de densidad de carga desdoblan la banda por la mitad, originando una banda completamente ocupada y otra completamente vacía, separadas por una brecha energética (*abajo*). Tal configuración no puede conducir una corriente sin dificultad, puesto que no es fácil producir el corrimiento en los niveles energéticos. Para que el sistema pueda conducir, debe suministrársele suficiente energía que induzca los electrones de valencia por encima de la brecha energética. (Dibujos de Gabor Kiss.)

gunos electrones de valencia por encima de la fosa energética; es decir, del rango de niveles energéticos prohibidos por las leyes de la mecánica cuántica. Cualquier movimiento neto de los electrones de valencia en tal material se halla inhibido, puesto que no existe lugar a donde aquéllos puedan ir dentro de los orbitales atómicos solapantes del material. Así, un material conduce fácilmente una corriente sólo si algunos de los estados energéticos en la banda de energía de los electrones de valencia no están ocupados por electrones.

En muchos materiales orgánicos, los orbitales de valencia ocupados en parte favorecen energéticamente su combinación química con otras sustancias. De aquí que el requerimiento de la ocupación parcial para la conducción eléctrica pueda llevar a la preparación de una sustancia químicamente inestable. Por ejemplo, el primer compuesto orgánico buen conductor fue descrito en 1954 por Hideo Akamatu, Hiroo Inokuchi y Yoshia Matsunaga, de la Universidad de Tokio. Se obtuvo ese compuesto tratando con bromo la molécula orgánica del perileno, pero se descomponía casi inmediatamente dando, por reacción ulterior con el propio bromo, una sustancia no conductora. Por eso, debido a tal inestabilidad química, sólo ha sido posible construir pocas moléculas aptas para la construcción de conductores orgánicos.

La estructura de cualquier auténtico conductor orgánico está asociada probablemente a dos condiciones. Primera, sus unidades estructurales moleculares deben acoplarse íntimamente a fin de que los electrones de conducción puedan moverse sin dificultad de una molécula a otra. Segunda, el coste energético para ocupar o desocupar parcialmente una banda de energía de valencia ha de ser bajo. Todas las moléculas investigadas hasta ahora tienen un esqueleto molecular grande y plano, formado por átomos de carbono y de hidrógeno y otros, como nitrógeno, azufre y selenio. Los electrones de valencia de tales moléculas se encuentran característicamente situados por encima y por debajo de su armazón plano. De ahí que las moléculas puedan apilarse apretadamente como en un pastel de pisos de tortillas y los electrones de valencia propagar una corriente de plano en plano por toda la pila.

Una de las primeras moléculas de ese género sintetizada fue el 7,7,8,8-tetraciano-*p*-quinodimetano, denominado abreviadamente TCNQ. Fue preparado en 1960 por investigadores de la E. I. du Pont de Nemours and Company. Se precisa poca energía para introducir un electrón adicional en la molécula del TCNQ y, por tanto, para iniciar la ocupación de una nueva banda energética de valencia en la pila de tales moléculas. La estructura resultante, cargada negativamente, es estable químicamente. Sin embargo, el propio TCNQ es incapaz de conducir la corriente eléctrica. En efecto, todas las moléculas de TCNQ poseen la misma afinidad por los electrones y, en consecuencia, en un sistema eléctricamente neutro no existe tendencia alguna a moverse los electrones de una molécula de TCNQ a otra. Si las moléculas TCNQ pueden adquirir electrones de otros átomos o moléculas de su entorno que los cedan fácilmente, entonces el material es conductor. En la sal de cesio del TCNQ (Cs_2TCNQ_3), por ejemplo, cada átomo de cesio cede un promedio de dos tercios de electrón a cada molécula de TCQN. Por consiguiente, dos de cada tres moléculas de TCNQ se cargan negativamente, y si se aplica un campo eléctrico, los electrones adicionales pueden pasar de las moléculas TCNQ cargadas a las neutras.

En los primeros años setenta, Fred Wudl, de la Universidad del Estado de Nueva York en Buffalo, preparó una molécula orgánica llamada tetratiaful-

valeno, o TTF, que contiene cuatro átomos de azufre. El TTF es un complemento del TCNQ, puesto que cede fácilmente un electrón, formando una estructura estable cargada positivamente. Cuando Wudl combinó el TTF con átomos de cloro, éstos aceptaron algunos electrones cedidos por el TTF; el producto resultó ser un material fuertemente conductor.

En 1973, poco después de que se sintetizara el TTF, Dwaine O. Cowan y John P. Ferraris, de la Universidad Johns Hopkins, y Alan J. Heeger y Anthony F. Garito, de la Universidad de Pennsylvania, combinaron independientemente ambas moléculas orgánicas formando el TTF-TCNQ. En la forma sólida de esta substancia, las moléculas de TTF y de TCNQ se hallan apiladas en columnas separadas, y los electrones son cedidos de la pila de TTF a la de TCNQ. Estudios de rayos X, llevados a cabo por Robert Comes y sus colaboradores en la Universidad de París-Sud, en Orsay, han demostrado que tal transferencia entre las dos pilas tiene lugar en la proporción de 0.59 electrones por molécula. Por ello, el material resulta ser conductor.

La arquitectura del cristal de TTF-TCNQ determina una propiedad eléctrica sorprendente: este material es fuertemente conductor en una dirección. En la dirección más favorecida resulta ser unas 500 veces más conductor que en la que lo es menos. La causa de esa conductividad anisotrópica estriba en que todas las moléculas constituyentes del cristal se apilan en planos paralelos. Los orbitales de las moléculas interaccionan principalmente a lo largo de las pilas.

En vista de las sugerencias de Little acerca de la superconductividad orgánica, resulta muy interesante el descubrimiento de tal conductor unidimensional. Bajo condiciones extremas de temperatura y presión, o con ligeras modificaciones de su estructura molecular, ¿podría lograrse tal vez que el TTF-TCNQ fuera un superconductor, de acuerdo con el mecanismo propuesto por Little? Según se desprende de cálculos teóricos, es imposible causar la formación de pares de Cooper en un conductor estrictamente unidimensional a cualquier temperatura superior al cero absoluto. Así pues, no puede existir un superconductor unidimensional. (No obstante, ello no refuta la formación de pares Cooper mediante electrones polarizables, que es la esencia de la teoría de Little.) Por añadidura, la estructura de las pilas de TTF-TCNQ es una realidad muy diferente de la estructura de las cadenas conductoras lineales descritas por Little.

A pesar de las diferencias, no deja de ser meritoria la analogía con el superconductor hipotético de Little. En cualquier sistema molecular real, los conductores son unidimensionales sólo aproximadamente; siempre existe alguna interacción entre las pilas moleculares que permitiría el salto de pares de Cooper de una pila a otra. De ahí que el obstáculo teórico para la superconductividad en sistemas unidimensionales no lo es en la práctica. Además, ciertas investigaciones sobre las propiedades del TTF-TCNQ han dado lugar al desarrollo de compuestos afines que son quasi-unidimensionales y superconductores. No se sabe aun si el mecanismo de apareamiento de Cooper en tales sistemas es precisamente el que Little había previsto.

A fin de crear la superconductividad en materiales orgánicos, los investigadores han tenido que estudiar y superar problemas inherentes a los sistemas unidimensionales. Una de las propiedades más importantes de tales sistemas había sido observada ya por Rudolf Peierls, de la Univesidad de Birmingham, en 1954. Peierls puntualizó que la separación uniforme entre los planos adyacentes de las moléculas, característica de substancias como el TTF-TCNQ a temperatura ambiente no está favorecida energéticamente a toda temperatura. A medida que ésta disminuye, la energía total del sólido puede reducirse si resulta modificado el espaciado y la distribución de la densidad de carga de los electrones de valencia.

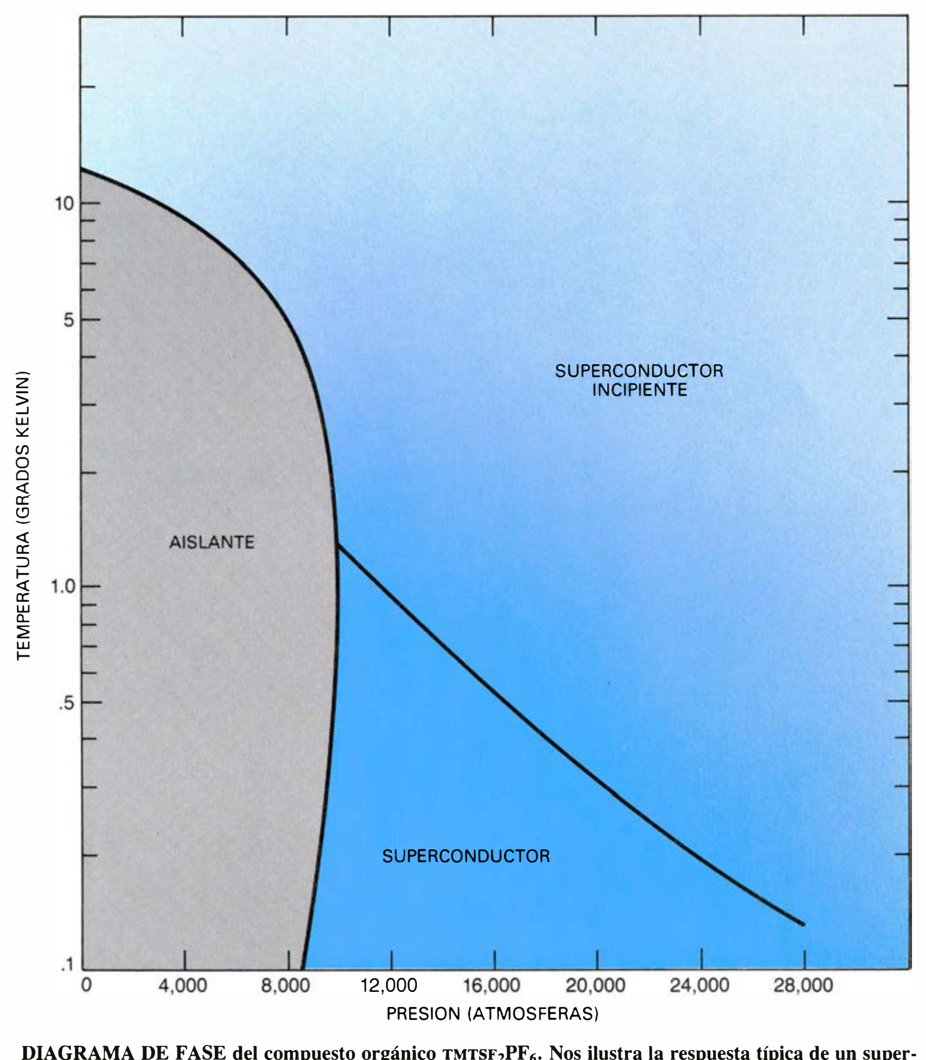
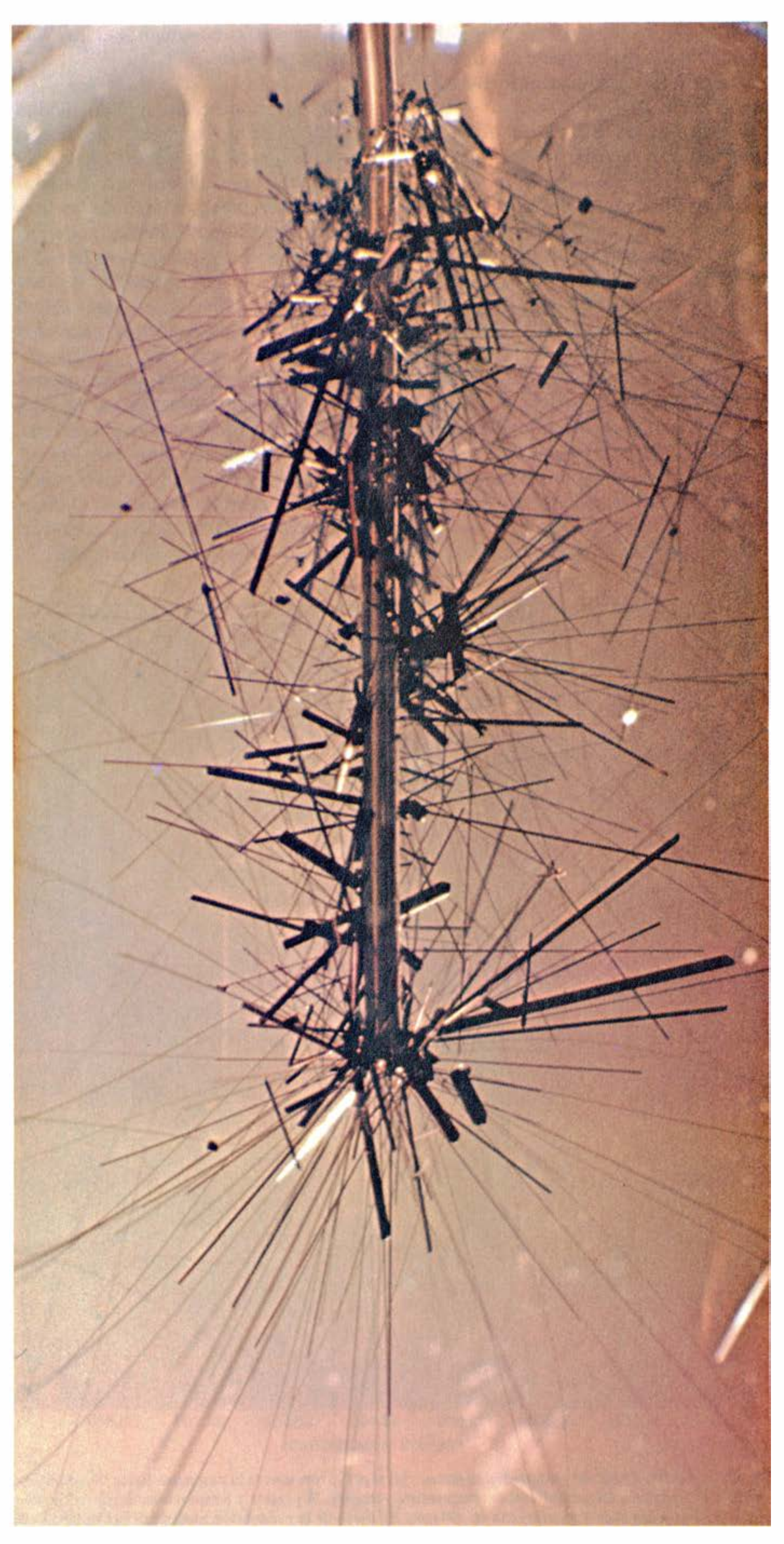


DIAGRAMA DE FASE del compuesto orgánico $TMTSF_2PF_6$. Nos ilustra la respuesta típica de un superconductor orgánico a las variaciones de temperatura y presión. A presión y temperaturas bajas (*zona gris*) el material es magnético y se convierte en aislante. El efecto de la ordenación magnética es similar al del establecimiento de una onda de densidad de carga; a saber, se crea una brecha entre las partes ocupadas y vacías de la banda energética de los electrones de valencia. A presiones mucho más elevadas, el material se vuelve superconductor (*región de color intenso*). Los autores de este artículo han propuesto que, fuera de estas dos zonas y hasta cerca de 30 grados Kelvin, la superconductividad es incipiente: los pares de Cooper se forman a lo largo de las pilas de moléculas durante períodos breves y distancias cortas.



SE FORMAN CRISTALES DE TMTSF$_2$PF$_6$ pasando una corriente eléctrica a través de una solución que contiene moléculas de TMTSF e iones negativos PF_6^-. Los electrones de valencia de las molécuals de TMTSF son atraídos por el electrodo positivo y en él se congregan en forma de cristales de pureza elevada. Los cristales aciculares que aparecen en esta fotografía poseen una longitud de cinco a 10 milímetros.

La densidad de carga de un electrón sufre concentraciones y rarefacciones denominadas ondas de densidad de carga. Cuando se establece la onda de densidad de carga y el retículo se distorsiona, la banda energética parcialmente ocupada por los electrones de valencia se desdobla en dos; la banda inferior queda completamente ocupada, y, completamente vaçía, la superior. Después de tal desdoblamiento, la conductividad puede establecerse sólo si mediante una fuente externa de energía térmica o electromagnética se promocionan electrones que superen la fosa energética existente entre ambas bandas.

Así pues, el establecimiento de una onda de densidad de carga puede hacer que un material pase bruscamente de conductor a semiconductor o, incluso, aislante. La transición suele tener lugar entre 50 y 100 grados K, y este mecanismo determina que el TTF-TCNQ se convierta en un semiconductor por debajo de 53 grados.

Los investigadores, a fin de modificar las propiedades dadoras del TTF y de otras moléculas similares, realizaron la substitución de los átomos de azufre por átomos de selenio, que son más pesados. Edward M. Engler, de la International Business Machines Corporation (IBM), preparó el tetraselenafulveno-TCNQ, o TSF-TCNQ, y uno de nosotros (Bechgaard) preparó el tetrametil-TSF, o TMTSF, cuando trabajaba con un grupo de investigadores de la Universidad Johns Hopkins. Esta última molécula se empleó finalmente para la obtención de cristales moleculares orgánicos conductores.

Inicialmente, el TMTSF parecía poco prometedor. Haciéndolo reaccionar con el TCNQ, se obtuvieron unos cristales rojos aislantes. El TMTSF-TCNQ es perfectamente estable en tal estado; con suerte y técnicas experimentales elaboradas, partiendo de él es posible preparar unos cristales negros y conductores, que constituyen una fase diferente correspondiente a otra configuración estructural. Por debajo de 60 grados K, una onda de densidad de carga desdobla la banda energética de valencia, y el material se convierte de nuevo en aislante.

En colaboración con Jan R. Andersen, del Risø National Laboratory de Roskilde en Dinamarca, hemos hecho reaccionar el TMTSF con una substancia llamada 2,5-dimetil-TCNQ, o DMTCNQ, estrechamente relacionada con el TCNQ; los cristales resultantes conducen la corriente a la temperatura ambiental a lo

largo de las pilas paralelas de moléculas, de manera similar a la observada en el TTF-TCNQ, pero este material se vuelve aislante por debajo de los 41 grados K. En Orsay, hemos empezado a estudiar los efectos de las altas presiones sobre las propiedades electrónicas del TTF-TCNQ y de sistemas análogos. En 1973, Meir Weger, de la Universidad Hebrea de Jerusalén, sugirió que la formación de la fase aislante de tales materiales, causada por las distorsiones de la densidad de carga, podría ser inhibida obligando a apiñarse sus moléculas por aplicación de presiones elevadas. Por consiguiente, nosotros esperábamos que una presión elevada podría inducir la superconductividad en el TTF-TCNQ, pero hallamos en cambio que las presiones elevadas ayudaban a estabilizar la fase aislante del material.

Sin embargo, si se aplica presión al TMTSF-DMTCNQ a la temperatura ambiente, se genera una nueva fase cristalina. Si se enfría el material bajo presión, no se vuelve aislante sino que, a temperatura del helio líquido, es decir, por debajo de cuatro grados K, continúa siendo superconductor.

Aun sin comprender del todo el fenómeno de la transición de fase, la estabilización del estado fuertemente conductor del TMTSF-DMTCNQ constituyó un hito en la búsqueda de un superconductor orgánico. A baja temperatura y a la presión de 10.000 atmósferas, la conductividad del material se sitúa por encima de los 200.000 mhos por centímetro, valor superior en un factor de 10 a la conductividad máxima del TTF-TCNQ. (El mho, la unidad de conductancia, es el recíproco del ohm, la unidad fundamental de resistencia.)

Aunque la conductividad de los mejores metales conductores, como el cobre y la plata, llega a ser unas 100 veces la del TMTSF-DMTCNQ, la movilidad de los electrones conductores a lo largo de las pilas de una molécula orgánica es mucho mayor que en un metal. Un electrón en la pila es dispersado por una de cada mil moléculas; en cambio en un electrón de un metal como el cobre esto sucede en uno de cada 50 a 100 espaciados del retículo, a menos que se trate de un metal excepcionalmente puro. (La mayor conductividad de un conductor metálico débese a un número de electrones mucho más elevado.) Se puede suprimir el estado fuertemente conductor del TMTSF-DMTCNQ aplicando un campo magnético perpendicular a las pilas o también sometiendo el material a una pequeña dosis de radiación.

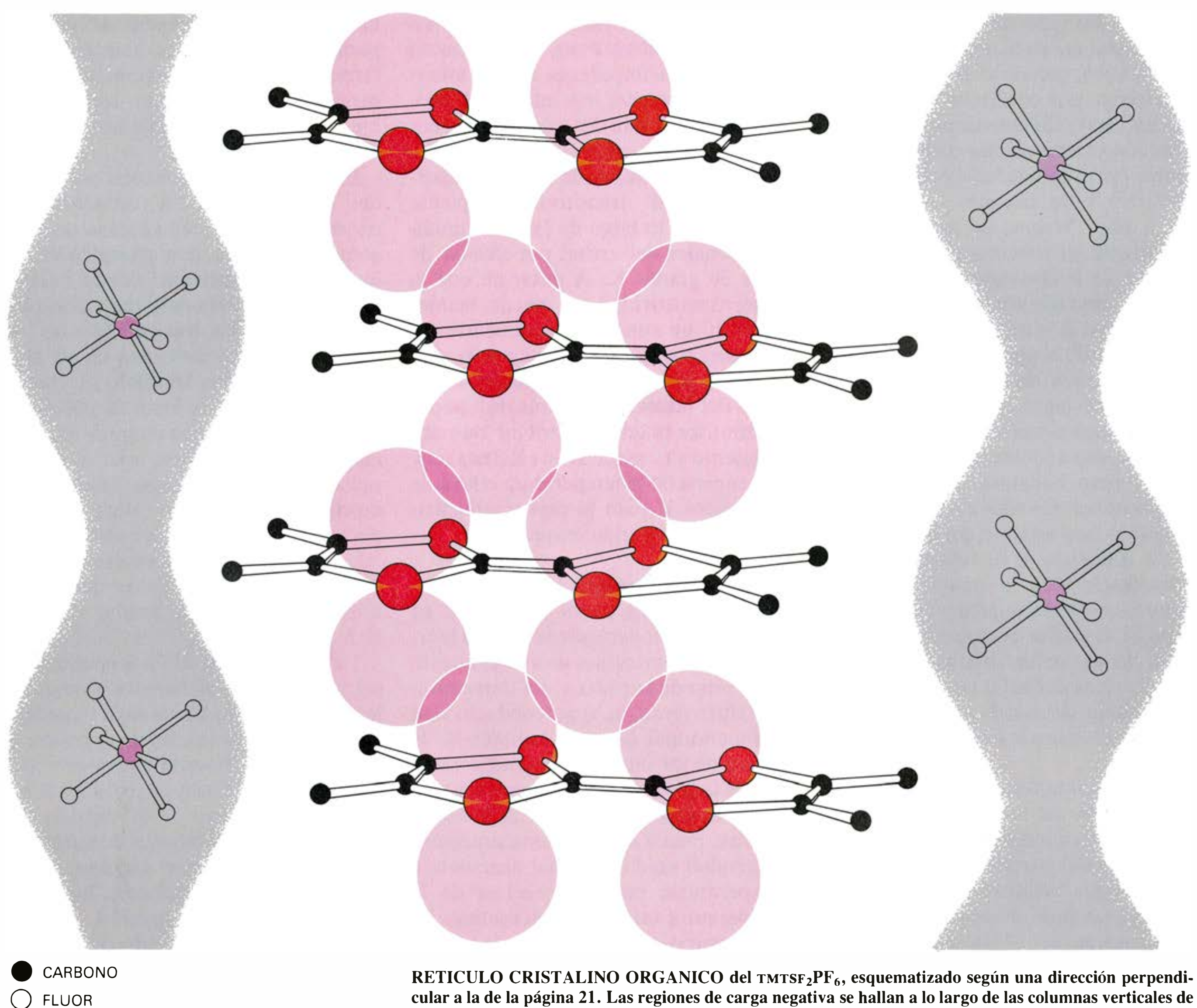
No se ha observado superconductividad en el TMTSF-DMTCNQ, tal vez a causa de pequeñas impurezas en las muestras. No obstante, sus intrigantes propiedades nos han llevado a proponer que su elevada conductividad es la manifestación o impronta de una superconductividad transitoria incipiente que surge a lo largo de las pilas unidimensionales del cristal por debajo de unos 30 grados K. A pesar de que la superconductividad no pueda mantenerse en un conductor tridimensional, el apareamiento de Cooper puede producirse a lo largo de una cadena molecular del mismo. Los pares de Cooper transitorios pueden contribuir substancialmente a la conducción eléctrica muy por encima de la temperatura crítica de 1,2 grados K para la superconductividad en masa o tridimensional.

La temperatura crítica para la formación de los pares de Cooper en una dimensión depende sólo de la fuerza de las interacciones de apareamiento en el seno de la cadena. La temperatura crítica para la superconductividad tridimensional depende también de la fuerza de las interacciones entre cadenas. Si la fuerza de las primeras interacciones es mucho mayor que la de las segundas, puede aparecer una superconductividad unidimensional incipiente a temperaturas muy por encima de la temperatura crítica tridimensional. Es de esperar que, a medida que la temperatura disminuya, vaya estableciéndose gradualmente la formación de pares de Cooper en una dimensión.

Nuestras investigaciones sobre el TMTSF-DMTCNQ han mostrado que las propiedades de este material están determinadas principalmente por la pila de moléculas de TMTSF. Por consiguiente, decidimos preparar un grupo de materiales en los cuales se exaltaran las propiedades de las pilas de TMTSF. A fin de lograr cristales de elevada pureza, disolvimos moléculas de TMTSF neutras en un disolvente orgánico; añadimos después una sal que diera iones cargados negativamente. Introdujimos luego en la solución dos electrodos de platino y pasamos una corriente eléctrica a través de ella. Los electrones de las moléculas de TMTSF desaparecieron en el electrodo positivo, dando moléculas cargadas positivamente. A medida que éstas se formaban, cada una se combinaba espontáneamente con una molécula neutra de TMTSF y con otra de los iones cargados negativamente, separándose el producto en el propio electrodo positivo en forma de una sal insoluble. De esta manera, se han preparado muchas de las llamadas sales $TMTSF_2X$, en donde X representa un ion cargado negativamente como, por ejemplo, el hexafluorofosfato (PF_6^-) o el perclorato (ClO_4^-).

Las sales $TMTSF_2X$ tienen propiedades excepcionales. A diferencia de otras pilas de moléculas orgánicas, tales como el TTF-TCNQ, la mayoría de las sales $TMTSF_2X$ continúan siendo fuertemente conductoras muy por debajo de 20 grados K. La transferencia de los electrones de las moléculas de TMTSF a las otras moléculas del sólido y la disposición regular de los iones cargados negativamente en el cristal impiden que el material sea susceptible a las distorsiones causadas por la onda de densidad de carga. A baja temperatura, un conductor orgánico que no es susceptible a tales distorsiones no se vuelve aislante, sino superconductor. La conductividad normal no persiste al enfriar el material.

En diciembre de 1979, sometimos la sal $TMTSF_2PF_6$ a condiciones similares a las que rindieron fuertemente conductor al TMTSF-DMTCNQ. En un experimento complicado, llevado a cabo por Michel Ribault en Orsay, sometimos la sal $TMTSF_2PF_6$ primero a la presión de 12.000 atmósferas y luego la enfriamos lentamente, a presión constante, durante un tiempo de 12 horas, hasta llegar por debajo de un grado K. Al mismo tiempo que se enfriaba el material se medía su conductividad y su susceptibilidad magnética. La resistencia eléctrica empezó a disminuir, y a 0,9 grados lo hizo bruscamente, anulándose. Las medidas magnéticas mostraron asimismo que a 0,9 grados la muestra expulsaba un débil campo magnético, lo cual es una indicación del efecto Meissner de los superconductores metálicos. Así pues, el material se había vuelto superconductor. Desde entonces, hemos observado la superconductividad en cristales de $TMTSF_2X$; siendo X hexafluoroarsenato (AsFe), hexafluoroantimoniato (SbFe), hexafluorotantalato (TaFe), perrenato (ReO_4^-) y ClO_4^-. Las propiedades de los cristales dependen algo de las que caracterizan a los iones cargados negativamente, aunque éstos no contribuyan directamente en el proceso de la conducción. A la presión atmosférica y a 12 grados K, la sal de PF_6^- se vuelve magnética y aislante y la sal de ReO_4^- da lugar a distorsiones de la onda de densidad de carga, convirtiéndose en aislante a 180 grados. Se precisan presiones elevadas para rendir superconductores a esas dos cla-



RETICULO CRISTALINO ORGANICO del TMTSF$_2$PF$_6$, esquematizado según una dirección perpendicular a la de la página 21. Las regiones de carga negativa se hallan a lo largo de las columnas verticales de iones PF_6^- (*sombreado gris*), mientras que aquellas otras de carga positiva se hallan a lo largo de las pilas centrales de las moléculas de TMTSF (*sombreado rojo*). Los orbitales de los electrones de valencia de las moléculas de TMTSF se solapan a lo largo de la pila. Por consiguiente, la conductividad del cristal en dirección paralela al eje de las pilas es mucho mayor que la conductividad en otras direcciones.

ses de cristales. Por otro lado, a 1,2 grados K y a la presión atmosférica, los cristales de TMTSF$_2$ClO$_4$ se vuelven superconductores.

Se dispone ya de un volumen considerable de datos experimentales que apoyan nuestra hipótesis, según la cual la superconductividad se halla en estado incipiente a lo largo de pilas moleculares unidimensionales a temperaturas muy por encima de la aparición de la superconductividad en masa. Los materiales absorben radiaciones a frecuencias del infrarrojo lejano por encima de un umbral de frecuencias que está relacionado con la energía de unión de los pares de Cooper en una cadena sencilla. Se han realizado estudios acerca del movimiento de los electrones desde un superconductor hasta una capa fina de un semiconductor aplicado sobre su superficie, que sugieren asimismo que los pares de Cooper se forman en cada cadena a temperaturas relativamente altas. Por último, puede reducirse la elevada conductividad de los materiales a temperaturas muy por encima de la temperatura crítica colocándolos en un campo magnético.

Estos y otros resultados indican que, en una cadena aislada, el valor de temperatura por debajo de la cual los electrones forman pares de Cooper tiende a ser el mismo; a saber, de unos 30 grados K. Así pues, las temperaturas críticas relativamente bajas, observadas en superconductores orgánicos, no son temperaturas límite superiores. Por el contrario, los experimentos realizados hasta ahora sugieren vivamente la existencia de una superconductividad transitoria y unidimensional a temperaturas superiores que las observadas en cualquier otro superconductor.

Hay todavía mucho que aprender acerca de la superconductividad de los materiales orgánicos. Por ejemplo, no se sabe si los pares de Cooper se forman a causa de las interacciones promovidas por el retículo, como sucede en los metales superconductores comunes, o bien por un mecanismo de mediación electrónica, como propuso Little. Puesto que los sistemas unidimensionales se conocen mejor, la química sintética puede abrir la senda que permita el desarrollo de materiales superconductores con aplicaciones tecnológicas.

Aviones ultraligeros

La unión del planeador de piloto suspendido y el pequeño motor ha dado lugar al "vehículo aéreo de recreo". Un aparato normal puede transportar 90 kilogramos y volar a 80 kilómetros por hora

Michael A. Markowski

Uno de los objetivos más perseguidos desde los primeros días de la aviación ha sido un avión económico que no presente excesivas dificultades para volar. Sin embargo, desde los tiempos de los hermanos Wright se ha tendido, casi siempre, hacia aviones más grandes y complicados. Sólo en los últimos años se ha materializado el tan ansiado objetivo en el avión ultraligero, unión de planeador de piloto suspendido y pequeño motor (de Kart o motodeslizador de nieve). Los numerosos modelos disponibles actualmente tienen precios que van de los 2800 a los 7000 dólares y, para un piloto que haya recibido la instrucción adecuada y que sea especialmente sensible a los caprichos del viento, el vehículo no es difícil de manejar.

Hasta que apareció en el mercado el avión ultraligero, la mayoría de los aviones se diseñaban y construían para satisfacer necesidades comerciales o militares. Los mismos aviones ligeros eran tan caros y costaban tanto de mantener que poca gente podía permitirse tener uno exclusivamente para recreo. El avión ultraligero es el primero que se ha desarrollado y vendido con éxito como "vehículo aéreo de recreo" (abreviado generalmente como VAR): el pasado año se vendieron más de 10.000 de estas aeronaves, superando el número de ventas de la aviación general (aviones utilizados con propósitos comerciales que no sean para los servicios regulares de pasajeros). Los ultraligeros empiezan a encontrar también ahora aplicación comercial en reconocimiento agrícola, fumigación y fotografía aérea. Algunos se han adaptado para servir en observación militar puesto que tienen la ventaja de no presentar imagen de radar.

Además de su disponibilidad, hay que considerar otros aspectos para entender la popularidad de los ultraligeros. Uno es el puro placer de volar de esta forma: el piloto de un ultraligero no va encapsulado en una cabina, sino que puede participar del viento. Otro, que con el motor parado el aparato puede volar a vela, como un planeador. Al final del día, el avión puede plegarse y guardarse en casa.

Más aún, la mayoría de los ultraligeros no tienen que registrarse ante los organismos competentes, ni se exige que sus pilotos tengan licencia. Hasta hace poco estas exenciones se admitían solamente para aparatos que pudieran despegar con el piloto soportando su peso, como las alas Rogallo. Sin embargo, al ir evolucionando los ultraligeros, los motores más grandes y los pesos más elevados han requerido la utilización de trenes de aterrizaje con ruedas. El otoño pasado la Administración Federal de Aviación de los Estados Unidos (FAA) inició un procedimiento destinado a establecer normas más claras sobre qué tipo de aeroplanos van a disfrutar de este privilegio. Si se aceptan las recomendaciones de los fabricantes y otros líderes de la industria, los criterios principales serán un peso máximo en vacío de 100 kilogramos y una carga alar máxima de 14,6 kilogramos por metro cuadrado. Con estas restricciones, las naves exentas de registro tendrían la suficiente solidez de estructura y aterrizarían a velocidades inferiores a 50 kilómetros por hora, condición necesaria para su fácil manejo.

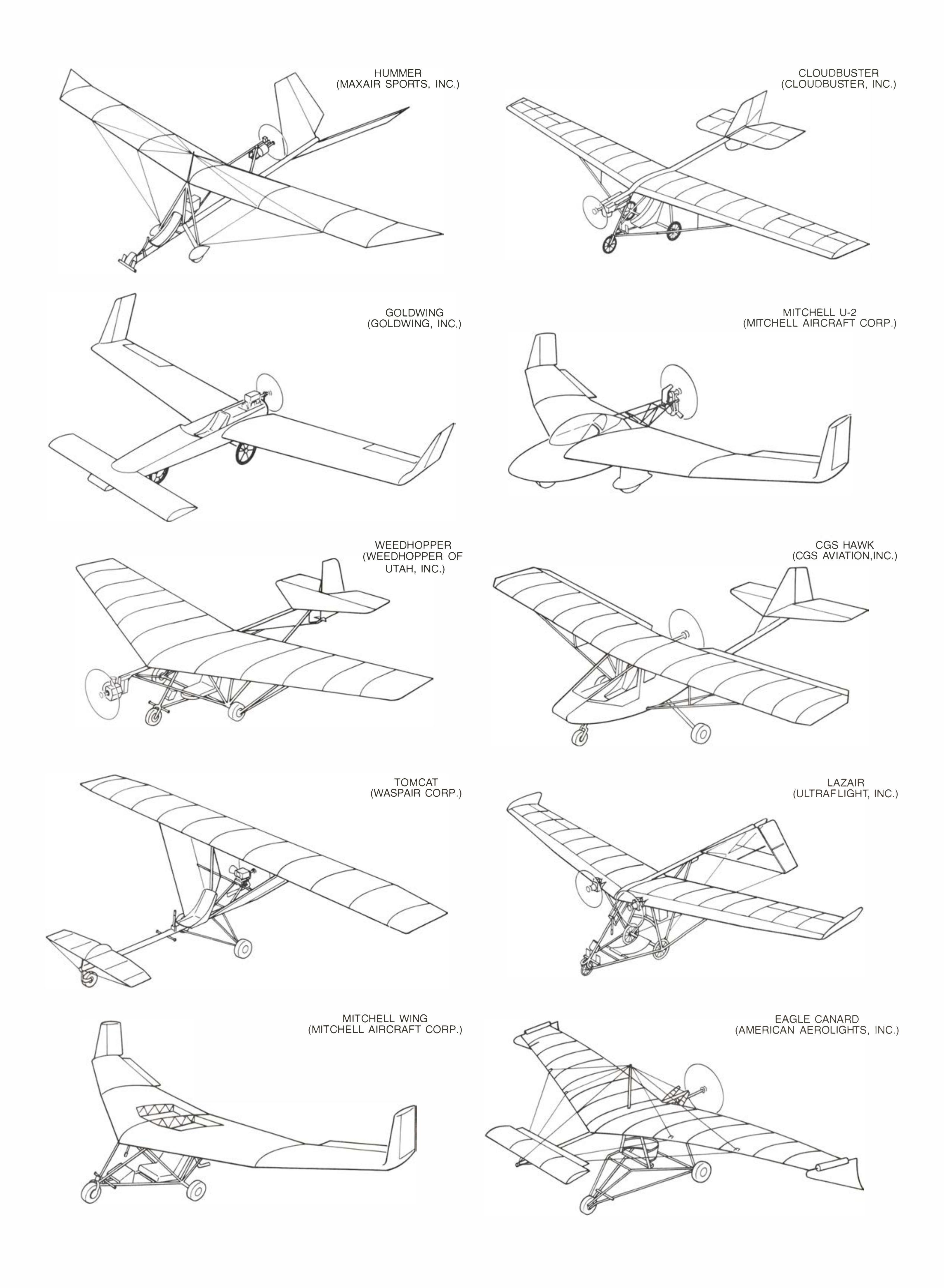
La mayoría de los modelos disponibles en la actualidad tienen un peso en vacío de unos 90 kilogramos; su envergadura es de 9 metros o más; la velocidad de crucero, de 80 kilómetros por hora, y la velocidad de pérdida de unos 40 kilómetros por hora. El coeficiente de planeo (la distancia recorrida por unidad de descenso) llega a 9:1 y la velocidad ascensional a 2,5 metros por segundo o más. La mayoría de estos vehículos pueden elevar más del doble de su peso en vacío, es decir, el piloto y el combustible pueden tener un peso conjunto de más de 90 kilogramos.

Los origenes del avión ultraligero se remontan a los primeros intentos de vuelo. El alemán Otto Lilienthal realizó más de 2000 vuelos durante la década de 1890, en lo que esencialmente eran planeadores de piloto suspendido. Lawrence Hargrave, de Australia, diseñó y ensayó aeromodelos propulsados mediante motor de gomas, de aire comprimido y de vapor. También inventó la cometa celular, que sirvió de base para todos los biplanos con montantes exteriores realizados posteriormente.

En los Estados Unidos, Octave Chanute, que en 1894 había escrito el clásico de la aeronáutica: *Progress in Flying Machines* (*Avances en máquinas voladoras*), diseñó un planeador basado en la cometa celular, incorporando un reticulado de Pratt, que en 1844 había sido patentado como método de reforzar puentes de ferrocarril. Las dos alas del planeador de Chanute estaban unidas por montantes verticales reforzados con cables cruzados en los planos longitudinal y transversal, formando una estructura rígida pero ligera que se convirtió en la elegida por los Wright y todos los diseñadores posteriores de biplanos. El planeador llevaba también un conjunto de cola posterior para la estabilidad y un perfil en curvatura para mejorar la sustentación.

El primero en volar un avión ultraligero motorizado fue probablemente Augustus M. Herring, que había sido ayudante de Chanute. Construyó una versión más pesada del planeador biplano de piloto suspendido de Chanute (en el que le había ayudado y había volado para él), montando un motor de aire comprimido de dos cilindros delante del ala inferior, que accionaba dos hélices de metro y medio dispuestas en tándem, una delante de las alas y otra detrás de ellas. El 11 de octubre de 1898, Herring realizó su primer vuelo

LA GRAN VARIEDAD de aviones ultraligeros hoy disponibles se apunta en las 10 versiones de la página siguiente. Pesan alrededor de 90 kilogramos y tienen una envergadura de 9 a 10 metros.


HUMMER
(MAXAIR SPORTS, INC.)
CLOUDBUSTER
(CLOUDBUSTER, INC.)
GOLDWING
(GOLDWING, INC.)
MITCHELL U-2
(MITCHELL AIRCRAFT CORP.)
WEEDHOPPER
(WEEDHOPPER OF
UTAH, INC.)
CGS HAWK
(CGS AVIATION,INC.)
TOMCAT
(WASPAIR CORP.)
LAZAIR
(ULTRAFLIGHT, INC.)
MITCHELL WING
(MITCHELL AIRCRAFT CORP.)
EAGLE CANARD
(AMERICAN AEROLIGHTS, INC.)

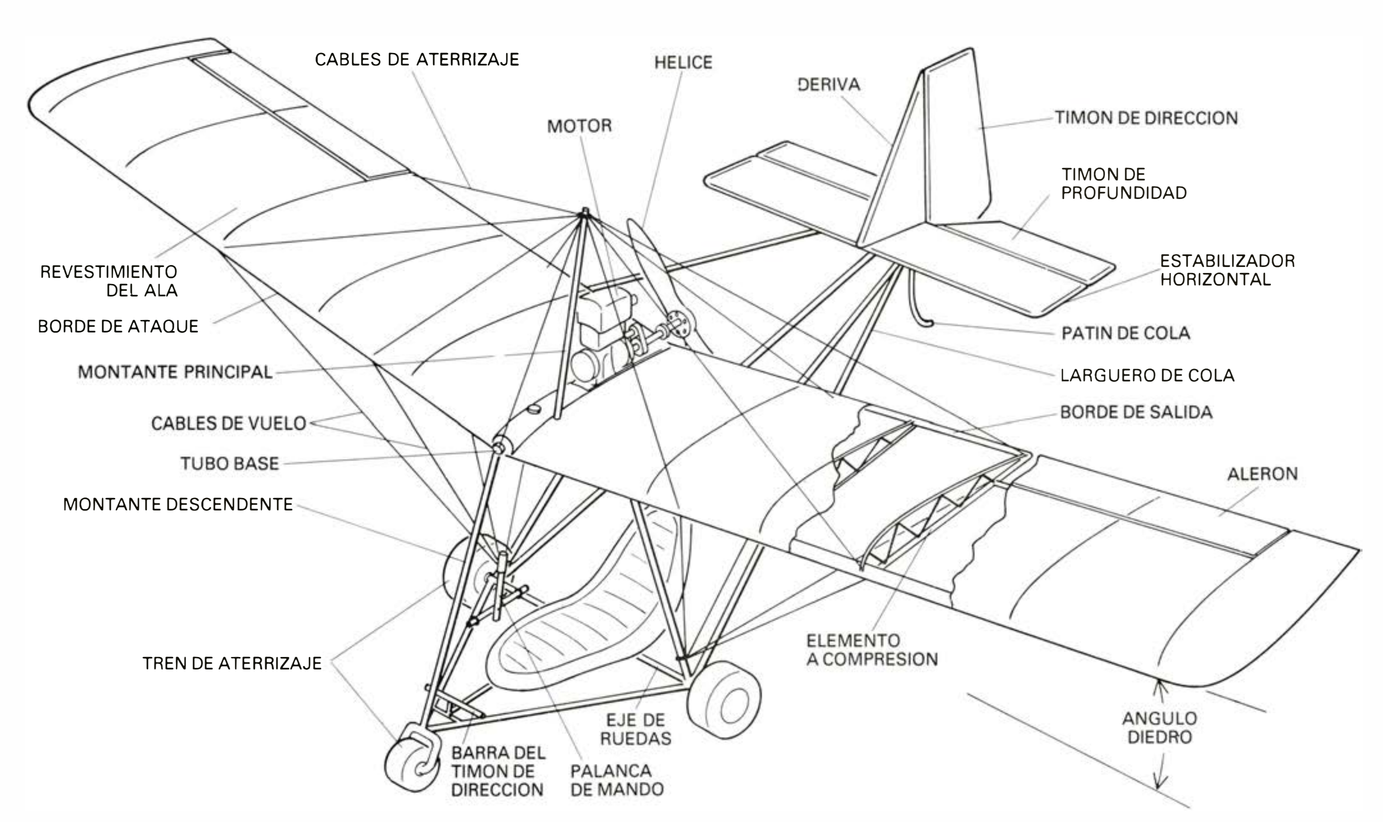
motorizado en St. Joseph, Michigan, recorriendo unos 81 metros en el aire, con un viento en contra de aproximadamente 40 kilómetros por hora. El único medio para controlar el planeador era desplazando su peso. Pocos días después realizó otro vuelo. Satisfecho porque había demostrado la posibilidad del vuelo motorizado, prosiguió con el diseño de un motor de vapor y un aeroplano mayor, que fueron destruidos por el fuego en 1899. Posteriormente organizó la Herring-Curtis Company, la primera firma de construcción de aviones de los Estados Unidos.

El éxito obtenido por los hermanos Wright al pilotar su avión el 17 de diciembre de 1903 supuso la consumación del primer vuelo motorizado en el que el piloto podía controlar el vehículo deflectando superficies aerodinámicas en lugar de desplazando su peso. El avión de los Wright fue un antecesor directo de los actuales aviones ultraligeros. En los años siguientes fueron apareciendo otros precursores: el Demoiselle, de Alberto Santos-Dumont, en 1909; el primer ultraligero verdaderamente profesional de D. W. Huntington, después de la Primera Guerra Mundial; el White Monoplane, en 1920; el inglés Electric Wren, en 1923; y el francés Pou-du-Ciel, en 1935. Ninguno de ellos tuvo éxito comercial apreciable.

El primer avión ultraligero moderno se fabricó en Wisconsin, en el invierno de 1974-75, y fue obra de John Moody, ingeniero eléctrico y piloto de vuelo libre. Montó un motor de Kart, de 12 caballos, en su biplano Icarus II de vuelo libre. Su único propósito era elevarse a altitudes donde poder manejar el vehículo como planeador, volando a vela y en planeo con el motor parado hasta que necesitara elevarse de nuevo. Sin embargo, no pasó mucho tiempo antes de que muchos de aquellos que siguieron su ejemplo empezaran a darse cuenta de que, añadiendo un motor a un planeador de piloto suspendido, se creaba un nuevo tipo de vehículo: un pequeño avión que podía volar independientemente de la sustentación natural requerida por el planeador de vuelo libre.

Pronto se adaptaron otros modelos de planeador para ser motorizados. Una prueba de este tipo demostró que hacer un avión ultraligero con éxito era algo más que montar un pequeño motor en un planeador. Fue en relación con el ala Rogallo, planeador de piloto suspendido adaptado a una cometa triangular, patentado por Francis M. Rogallo y su esposa en 1951. Al principio, el motor se fijó al montante principal del ala Rogallo adaptada para ser motorizada. La elevada línea de empuje a que daba lugar esta posición demostró ser desastrosa en los momentos en que la fuerza de gravedad no actuaba sobre el piloto. En esta situación de gravedad cero, una línea de empuje elevada creaba un momento de picado que el piloto no podía neutralizar. (En un vuelo normal, cuando el ala está bajo una carga *g* positiva, un desplazamiento del peso del piloto es efectivo como control, pero en gravedad cero, el piloto, de hecho, no tiene peso.) Con el morro hacia abajo, el revestimiento comenzaba a flamear, y el planeador quedaba obligado o bien a un picado que no podía pararse, o bien a un rizo invertido que probablemente causaría fallos estructurales. Afortunadamente, este peligroso defecto de diseño fue corregido con una línea de empuje más baja antes de que el ala Rogallo motorizada saliera al mercado.

La instalación más corriente del motor en el ala Rogallo consistía en una hélice impulsora en el extremo trasero de la quilla, que es el larguero central del ala. El equipo constaba de un motor de Kart montado sobre la cabeza del piloto y un largo eje de transmisión situado paralelamente a la quilla. Este mismo equipo podía montarse prácticamente en cualquier planeador de ala flexible. Un planeador así equipado

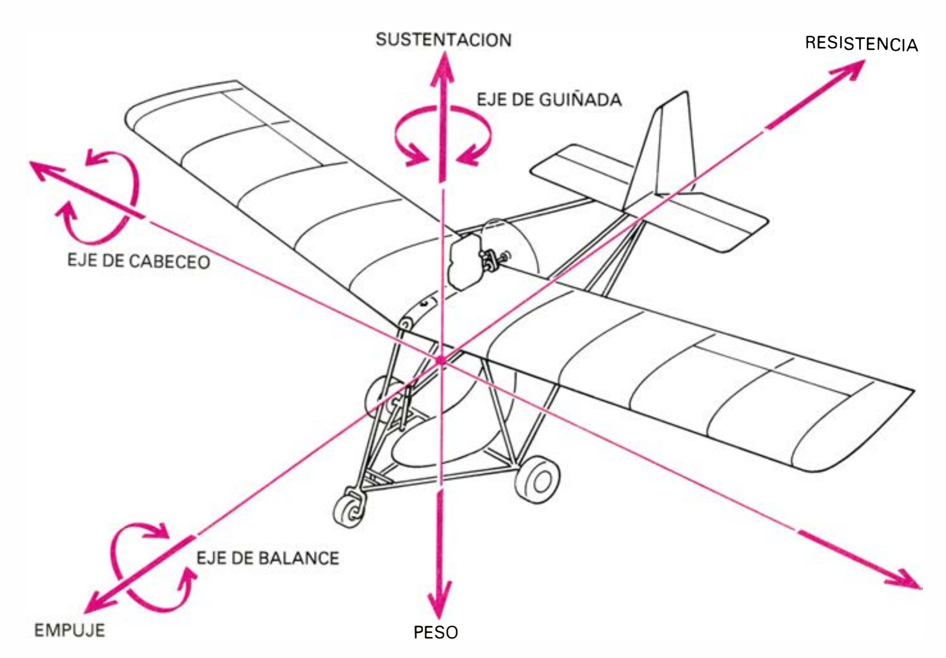


ELEMENTOS COMPONENTES de un avión ultraligero de tipo general. El vehículo representa la fase más reciente de diseño, pues está dotado de un sistema de control de tres ejes independientes: el cabeceo (movimiento del morro arriba y abajo) se regula moviendo la palanca de control hacia atrás y hacia adelante; el balance, moviéndola de lado a lado (haciendo que un alerón suba y el otro baje); y la guiñada, maniobrando el pedal de dirección con el pie. Un avión ultraligero con este sistema de control, que ya utilizaron los hermanos Wright, se pilota casi de la misma forma que un avión más pesado.

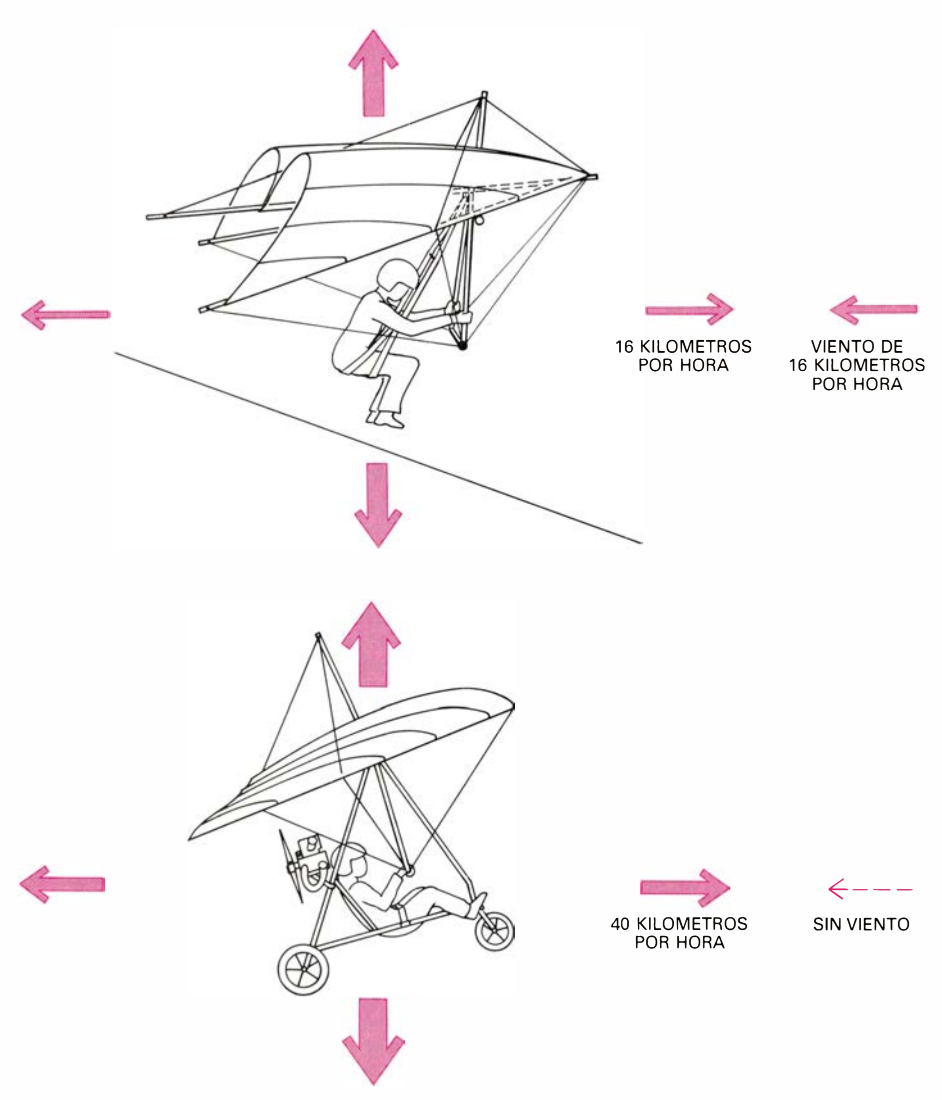
despegaba a pie, es decir, el piloto corría contra el viento con el equipo suspendido sobre su cabeza. Puesto que la línea de empuje aún estaba más alta que el centro de gravedad, no era recomendable volar en condiciones de turbulencia, porque con el motor en marcha el vehículo tendía a hacerse inestable longitudinalmente. Sin embargo, en condiciones de calma, esta situación era practicable. Este equipo se empleó más para ganar altura que permitiese el vuelo a vela que para el vuelo motopropulsado propiamente dicho.

Hasta 1977, la mayoría de los aviones ultraligeros obtenían el empuje mediante un motor de Kart, que accionaba una hélice montada directamente sobre el eje del motor. El sistema era sencillo y sin problemas mecánicos, pero el motor giraba a tal número de revoluciones por minuto que la hélice tenía que ser muy pequeña (de unos 71 centímetros de diámetro) para que sus puntas no superaran la velocidad del sonido. En las condiciones de funcionamiento, la hélice giraba a más de 9000 rpm, muy cerca de la velocidad del sonido, resultando que el nivel de ruido producido era elevado y el rendimiento propulsivo se situaba solamente alrededor del 50 por ciento. En consecuencia, el rendimiento de los primeros planeadores motorizados de piloto suspendido era completamente marginal. La velocidad de crucero apenas superaba a la de pérdida, y la velocidad ascensional era peligrosamente baja: 0,5 metros por segundo. Además, los motores mismos tenían una vida corta, pues debían girar a velocidades muy elevadas para generar suficiente potencia. Parecía evidente que los aviones ultraligeros no tendrían buena acogida hasta que se mejorase su nivel de actuaciones.

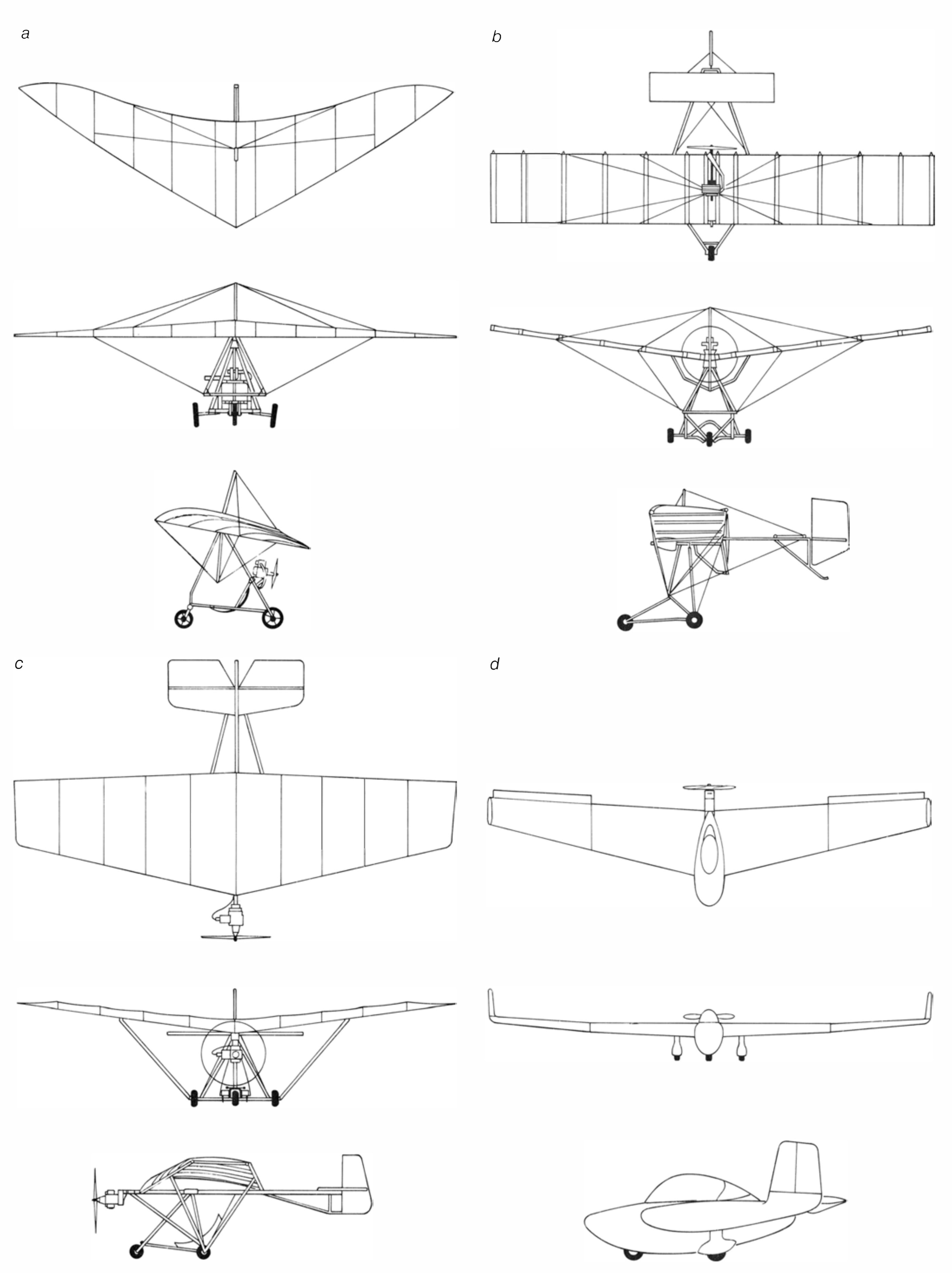
Charles Slusarczyk, un emprendedor investigador, abordó el problema del empuje de manera científica. Consciente de que una hélice es más eficiente cuando sus extremos se mueven bastante por debajo de la velocidad del sonido, ideó un sistema de transmisión reductora para los planeadores motorizados de piloto suspendido, que patentó el año pasado. Su idea consistía en mover un gran volumen de aire tan despacio como fuera posible a través del disco formado por la hélice al girar, minimizando al mismo tiempo la resistencia debida a la comprensibilidad, que absorbe gran cantidad de potencia a velocidades próximas a la del sonido. El resultado es una gran mejora en el empuje y en el rendimiento, y una disminución sustancial del ruido de la hélice. Casi todos los aviones ultrali-



CUATRO FUERZAS actúan sobre un avión en vuelo: la sustentación, el peso, el empuje y la resistencia. La sustentación la proporciona la diferencia de presiones del aire alrededor del ala, producida por el flujo de aire sobre ella. El peso es efecto de la gravedad. El empuje lo da la hélice. La resistencia es la resultante de las distintas fuerzas que frenan el avión. También se indican los ejes de cabeceo, balance y guiñada.



FUERZAS QUE ACTUAN EN EL DESPEGUE en un planeador de piloto suspendido de tipo Rogallo (*arriba*) y en un avión ultraligero (*abajo*), también de ala Rogallo. El piloto de planeador despega creando una sustentación en el ala, sobre su cabeza, al correr cuesta abajo en contra de un viento rápido. La sustentación y el empuje son mucho más débiles en el planeador de piloto suspendido que en el avión.



CUATRO TIPOS FUNDAMENTALES de avión ultraligero: el Rogallo (*a*), representado por el Jet Wing; el arriostrado mediante cables (*b*), representado por el Quicksilver, que es el avión que aparece en la portada de este número; el de montantes (*c*), aquí el Weedhopper; y el de ala en voladizo (*d*), que tiene un ala a cada lado del fuselaje soportadas únicamente por su unión a éste. El vehículo en voladizo que aquí vemos es un Mitchell U-2. El nivel de actuación de los aviones, en especial en lo que se refiere a su velocidad máxima, se refleja aproximadamente en la ordenación de las cuatro categorías.

geros actuales están dotados de un sistema de propulsión de transmisión reductora, que obtiene brillantes prestaciones de un motor pequeño. Un valor típico para las transmisiones reductoras actuales es de 4,5 kilogramos de empuje por caballo.

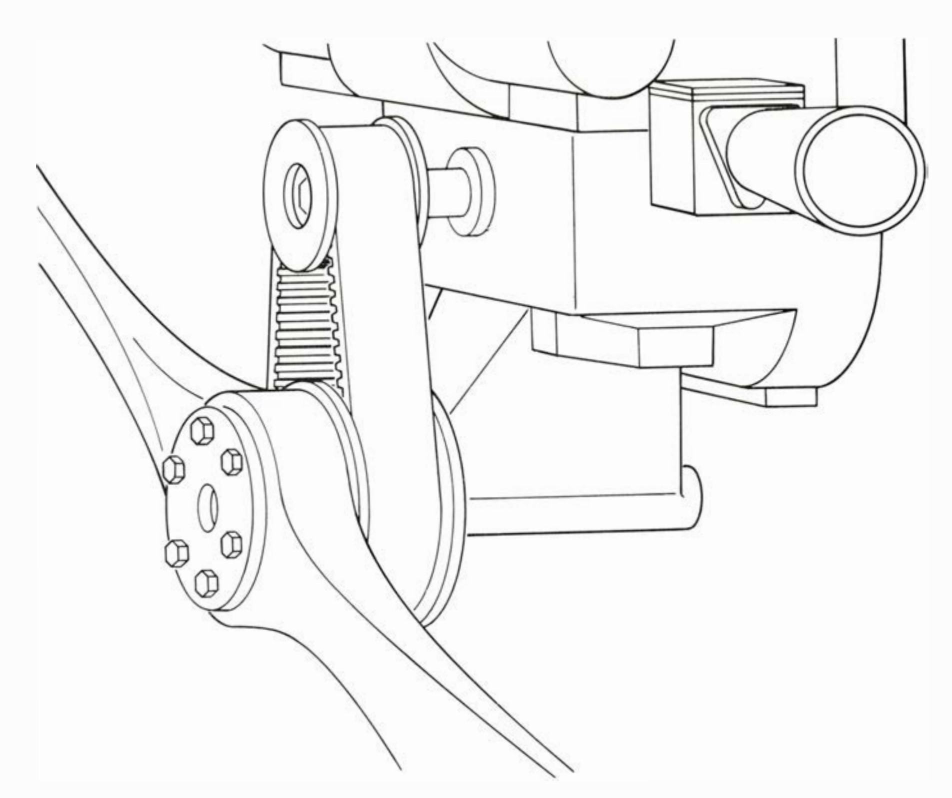
Poco después, el motor de Kart fue ampliamente sustituido por el de motodeslizador de nieve, que tiene un recorrido más largo y gira a menos revoluciones por minuto. La puesta a punto del motor se ajustó a las especificaciones de su nueva utilización, lo que proporcionó más empuje, mayor seguridad y un aumento de la esperanza de vida útil. Generalmente se reducía la carburación, o se relajaba la compresión, aunque a veces se practicaban ambas correciones.

Uno de los cambios que ha forzado el uso de motores de motodeslizador ha sido el aterrizaje con ruedas. El planeador de transmisión directa, con motor de Kart, tenía un peso en vacío de 45 kilogramos y podía despegar a pie. El motor de motodeslizador, junto con los componentes correspondientes a la transmisión reductora y el robustecimiento del armazón, han elevado el peso en vacío a más de 70 kilogramos, haciendo peligroso este tipo de despegue.

La incorporación del tren de aterrizaje conformó definitivamente el avión ultraligero. Una vez aceptado el tren de aterrizaje, los ingenieros podían ya diseñar vehículos ultraligeros partiendo de un sistema de propulsión, en lugar de limitarse a montar un motor en un planeador. El resultado ha sido una nueva generación de modelos, la mayoría de los cuales alejan esta nueva aviación de sus orígenes de planeador motopropulsado de piloto suspendido, aproximándola más al "avión de bolsillo".

Los aviones ultraligeros actuales se agrupan en cuatro categorías, que reflejan la estructura básica de cada diseño: Rogallo, de tirantes (cable braced), de montantes (strut braced) y en voladizo (cantilevered). El orden en que se presentan da una idea del nivel relativo de prestaciones de cada una de ellas, en especial respecto a su velocidad máxima.

La clase Rogallo no sólo incluye las versiones de planeadores de despegue a pie con un motor montado, sino también las naves conocidas como "trike". Un trike (de triciclo –tricycle– en inglés) es una estructura tubular piramidal que sostiene el motor, el asiento del piloto y el tren de aterrizaje. Todo el conjunto está unido a un planeador Rogallo de piloto suspendido. Como consecuencia de esta disposición, su propietario tiene, de hecho, dos aeronaves: un planeador de piloto suspendido y un avión ultraligero. En ambas, la forma de volar es la misma: el piloto empuja o tira de una palanca de control de cabeceo y desplaza su peso a un lado u otro para el control lateral.



TRANSMISION REDUCTORA, patentada el año pasado por Charles Slusarczyk, que mejora sustancialmente el empuje y el rendimiento de las hélices en los aviones ultraligeros. Una correa dentada transmite la potencia desde una pequeña polea del eje del motor a otra mayor, en el eje de la hélice. Una transmisión reductora típica de este tipo proporciona 4,5 kilogramos de empuje por caballo de vapor. Con la transmisión directa, las puntas de la hélice giran casi a la velocidad del sonido, algo poco eficiente.

Un avión ultraligero corriente, como el Eagle, incorpora un ala híbrida de Rogallo y un canard. El ala principal es un ala Rogallo modificada mediante la inserción de costillas y la incorporación de bordes marginales inclinados, o timones de dirección. El canard es una superficie rígida con un timón de profundidad que sirve para el mando de cabeceo. Dicho mando se incrementa mediante desplazamientos hacia adelante y hacia atrás del piloto.

La mayoría de los aviones ultraligeros actuales pertenecen a la categoría de los arriostrados mediante cables. Un sistema típico consiste en una estructura tubular piramidal que rodea al piloto, con un tren triciclo de aterrizaje sujeto a ella. El ala se une a la parte superior de la pirámide mediante un pasador y se arriostra con cable a la parte inferior de ésta y al montante principal por encima del ala. Otro larguero va hasta la parte posterior y sostiene la cola, que también está arriostrada por cables a la pirámide, al montante principal y al ala.

Generalmente el motor se monta en la sección central del ala. Una transmisión reductora lleva la potencia hasta una hélice impulsora. La estructura del ala se reviste con fundas de Dacrón para crear la superficie que proporciona la sustentación. En el extradós del ala se utilizan costillas de aluminio preconformadas que se insertan en bolsas del revestimiento a fin de crear un perfil con curvatura, lo que mejora la sustentación. El intradós es plano o ligeramente abombado para compensar el hecho de que en vuelo el aire tiende a empujar dicha superficie hacia arriba. Algunos modelos tienen revestimiento de Dacrón sólo en el extradós. Estos vehículos son más lentos que los de ala con doble superficie.

La construcción con montantes va ganando popularidad a medida que los diseñadores tratan de mejorar la aerodinámica del avión ultraligero. El grupo incluye una amplia variedad de estructuras alares. En los modelos más sencillos, la estructura alar se cubre con fundas de Dacrón y las costillas se conectan a la superficie superior. Los modelos más modernos incorporan un larguero, fabricado en tubo de aluminio, de sección en *D*, que sirve también como borde de ataque del ala. Los montantes se fabrican de aluminio o de un compuesto con núcleo de espuma y revestimiento de aluminio o fibra de

vidrio. Los fuselajes van de la simple pirámide tubular a las estructuras semicerradas fabricadas con componentes de espuma y fibra de vidrio.

Los aviones del cuarto grupo están construidos casi de la misma forma, pero tienen alas separadas en voladizo, es decir, un ala a cada lado del fuselaje, soportadas únicamente por su sujeción a éste. El modelo en voladizo da lugar a las naves ultraligeras más aerodinámicas, siendo también las más parecidas a los aviones convencionales. Otra de las virtudes del ultraligero en voladizo es que, una vez situado en el lugar de despegue, se tarda menos en montarlo.

Los controles de los aviones ultraligeros también han evolucionado significativamente en los últimos años. Cuando estos vehículos se fabricaban ajustando un motor a un planeador de piloto suspendido, el control solía consistir nada más que en el desplazamiento del peso del piloto. Muy pocas naves tenían algún control aerodinámico rudimentario, esto es, superficies que el piloto pudiera mover para variar el asiento, la guiñada o el balance. Los pilotos de planeador de piloto suspendido, familiarizados con estos sistemas de control, los consideraban adecuados, pero a los entrenados en aviones convencionales dichos sistemas les resultaban extraños y confusos. Los diseñadores de aviones ultraligeros se dieron cuenta pronto del problema y empezaron a desarrollar controles aerodinámicos adecuados. Hoy es raro encontrar un avión que dependa del desplazamiento del peso del piloto para su control.

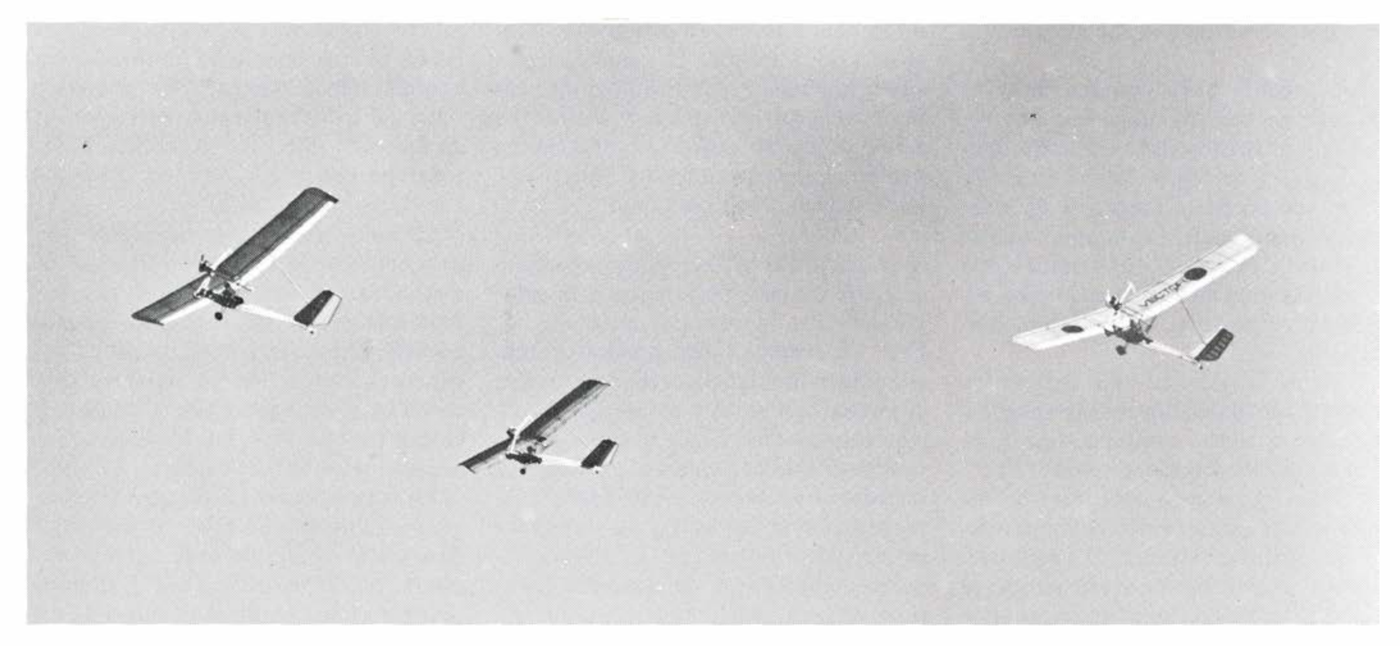
El primer avión ultraligero moderno, el planeador motorizado Icarus II, tenía un sistema de control híbrido, puesto que así era como se controlaban los pianeadores de piloto suspendido antes de que Moody pensara añadirle un motor. El piloto, que iba suspendido de un arnés en posición prono, controlaba el asiento y, por lo tanto, la velocidad, moviéndose hacia adelante o hacia atrás. También podía mover los bordes marginales para hacer que el vehículo guiñara, induciendo así un balance. En el giro, las alas interiores perdían velocidad y, por tanto, sustentación, mientras que las exteriores ganaban velocidad y sustentación, creando un momento de balance neto y un giro. La deflexión simultánea de los dos bordes marginales aumentaba la resistencia del avión, sirviendo así como medio de control de la senda de planeo.

Otro sistema híbrido era el empleado por el Quicksilver, que apareció originalmente como planeador de piloto suspendido a principios de los 70. El piloto iba suspendido en balancín, con lo que podía desplazar su peso de delante atrás y de un lado a otro. Para aumentar la capacidad de giro de la nave, se conectaban cables desde el arnés al timón de dirección, de forma que un desplazamiento lateral del piloto deflectaba el timón, provocando una guiñada y, por tanto, un balance. El sistema era práctico, pero seguía siendo del tipo de desplazamiento del peso, y los pilotos con licencia no lo aceptaron.

El primer avance importante en los sistemas de control apareció en el primer auténtico avión ultraligero (para diferenciarlo de los planeadores motorizados de piloto suspendido). Esta nueva idea consistía en un sistema de control alrededor de dos ejes. El piloto queda sujeto al asiento por un cinturón, no pudiendo desplazar su peso efectivamente. En su lugar, empuña una palanca de control que está conectada al timón de dirección y al correspondiente timón de profundidad. Las alas no tienen superficies móviles. La palanca de control funciona en parte de manera convencional: un movimiento hacia adelante introduce un picado; un movimiento hacia atrás, un encabritado. Los movimientos laterales de la palanca, sin embargo, mueven el timón en la misma dirección de la palanca, creando una guiñada y, por lo tanto, un balance. (En un avión con sistema convencional de control de palanca, un movimiento lateral de ésta mueve los alerones de las alas para conseguir el balance.)

En el sistema bicontrol, una deflexión del timón hace que el vehículo guiñe y deslice. La velocidad del ala exterior aumenta creando un momento de balance y su ángulo correspondiente. Por esta razón, las alas deben tener el suficiente diedro (ángulo que forman entre sí) para detener el deslizamiento conforme se produce. En una aeronave adecuadamente diseñada, el movimiento resultante es un giro razonablemente coordinado. El mayor inconveniente del sistema bicontrol es que el vehículo se hace difícil de manejar con viento cruzado superior a 8 kilómetros por hora, porque entonces las alas no se pueden levantar lo bastante rápido para evitar el deslizamiento.

Los modelos más recientes llevan un



CIRCO VOLANTE, en una reciente demostración de aviones ultraligeros en Lakeland, Florida. Consta de tres ultraligeros Vector 610. El 610 es un avión del tipo de arriostrado con cables, con una hélice impulsora y una cola vertical en *V*. El piloto va en una cabina bajo los 10,3 metros de ala del vehículo.

sistema de control alrededor de tres ejes independientes, que fue la clave del éxito de los hermanos Wright. Una versión de este sistema, que es esencialmente la misma de un equipo convencional, comprende el control del timón de profundidad y los alerones, por medio de la palanca, y el del timón de dirección, con pedales. Otra versión modificada utiliza spoilers (disminuidores) en lugar de alerones. Se coloca un spoiler (disminuidor) sobre la superficie de cada ala: deflectando uno de los spoilers, se rompe el flujo de aire por encima del ala correspondiente, reduciendo la sustentación y haciendo que el ala caiga. El spoiler deflectado incrementa también la resistencia, aumentando la guiñada e induciendo una sustentación adicional en el ala exterior debido a que aumenta su velocidad. Deflectando los dos spoilers simultáneamente el piloto puede controlar la senda de planeo.

Quienquiera que tenga la intención de volar un avión ultraligero debe dominar la aerodinámica básica. Las cuatro fuerzas que actúan sobre el vehículo son: la sustentación, el peso, el empuje y la resistencia. La sustentación la proporciona la presión del aire por debajo del ala y el flujo de aire sobre ella. El peso (del avión, el piloto y el combustible) es efecto de la gravedad. El empuje es la fuerza suministrada por la hélice. La resistencia es la resultante de fuerzas que tienden a frenar el avión. La resistencia inducida es un subproducto de la sustentación y aumenta al incrementar el ángulo de ataque (el ángulo entre la cuerda del ala y la dirección del viento relativo o, menos rigurosamente, entre el ala y la horizontal). La resistencia parásita la generan los elementos del avión que no contribuyen a la sustentación; la resistencia de perfil la genera el perfil del ala (forma de la sección recta del ala).

Lo más importante para el piloto es el ángulo de ataque. Cuando es grande, la velocidad de vuelo es baja, y cuando es pequeño, la velocidad de vuelo es alta. Aumentando el ángulo de ataque, se llega a un punto en el que el ala entra en pérdida, disminuyendo su sustentación y aumentando su resistencia, ya que el flujo de aire se hace completamente turbulento y se desprende de la superficie superior del ala. En los puntos próximos a la pérdida, la capacidad del piloto para controlar la nave se ve disminuida puesto que los alerones se vuelven prácticamente inútiles. Sólo una disminución del ángulo de ataque permitirá que el ala vuelva a ganar sustentación, lo que significa que el piloto debe empujar la palanca de mando hacia adelante para que el morro baje y el avión gane velocidad perdiendo altura. La pérdida puede convertirse en barrena, sobre todo en un viraje: si una de las alas entra en pérdida antes, ésta cae y el avión empieza a dar vueltas alrededor de un eje vertical mientras cae. No siempre es posible recuperarse de una barrena, por lo que es importante para el piloto (especialmente para los noveles) conocer las condiciones que pueden llevar a la pérdida, para evitarlas si es posible, aplicando las medidas de corrección adecuadas rápidamente si se entra en ella o parece inminente.

Aunque, básicamente, los aviones ultraligeros son más fáciles de manejar que los aviones más pesados y en los Estados Unidos no se exige licencia para pilotar la mayoría de ellos, la FAA exige que cualquiera que vuele esté familiarizado con las Normas Federales del Aire (Federal Air Regulations). Todos los pilotos de aviones ultraligeros deben prestar atención particular a las normas de vuelo visual en vigor. Debe evitarse el vuelo cerca de aeropuertos con control aéreo.

Sería temerario para cualquiera intentar volar un avión ultraligero sin adquirir primero un sólido conocimiento de aerodinámica y realizar algún vuelo de entrenamiento reglamentario. El ultraligero es un avión de verdad, no un juguete, y se necesita una buena dosis de pericia y conocimientos para pilotarlo con seguridad. Muchas escuelas de vuelo exigen a sus alumnos una previa instrucción en doble mando en un avión convencional y capacidad para hacer un vuelo solitario en él. Puesto que el ultraligero es un avión individual, parece prudente una previa enseñanza en doble mando para todos los que no sean ya pilotos. Deben conocerse los fundamentos del vuelo antes de aventurarse al vuelo en solitario.

Mencionaba al principio que el piloto de un ultraligero debe ser especialmente sensible a los caprichos del viento. Esta necesidad queda reflejada en el hecho de que los pilotos con licencia, que constituyen aproximadamente la mitad de los que vuelan vehículos ultraligeros, tienen a menudo más problemas que los no pilotos para adaptarse a estas condiciones: acostumbrados a aviones más pesados, parecen no tener suficientemente en cuenta al viento, siendo sorprendidos cuando el avión ultraligero pone de manifiesto su extremada sensibilidad a las ráfagas. Un piloto prudente sólo vuela un avión ultraligero cuando el tiempo es bueno y el viento suave y uniforme.

Ciencia y sociedad

Termodinámica y ecología

Sea un ecosistema planctónico. La radiación solar de las capas de agua iluminadas proporciona energía reductora a los cloroplastos de los productores primarios; en el interior de las células aparecen compuestos reducidos de los elementos biogénicos (C, N, S). Las estructuras materiales del organismo se hunden en el agua y aportan energía reductora hacia las profundidades, donde se liberan la materia orgánica y materiales reducidos (amoníaco, metano, sulfhídrico). La combinación de la luz y la gravedad definen un eje principal de organización, casi vertical, que presenta, entre otras propiedades químicas y bioquímicas, un gradiente de potencial de oxidación-reducción y un flujo de electrones en el agua opuesto al flujo neto de materia orgánica.

También en los ecosistemas terrestres, la luz y la gravedad definen y configuran un eje vertical en la organización. Un árbol es comparable con un ecosistema planctónico. Las unidades de producción, en este caso las hojas, se mantienen en un conjunto y están unidas a un sistema de transporte que aportan los organismos.

Cuando un ecosistema trabaja bajo una presión o stress foráneo, las respuestas que se generan son de dos tipos: los elementos recorren sus ciclos más deprisa y una fracción de los elementos incluidos en el ciclo quedan fuera de él, o se descargan en los límites, o a lo largo de trayectorias externas, a veces inmovilizados en el interior del mismo sistema. Los lagos eutróficos, los sistemas de surgencias marinas y los propios sistemas agrícolas sometidos a tensión ceden oxígeno y nitrógeno a la atmósfera y carbono orgánico y fosfato al sedimento o al suelo. Este mismo mecanismo ha funcionado con eficacia a lo largo de la historia entera de la vida, permitiendo la formación de la atmósfera terrestre a partir de la primitiva biosfera.

La luz y la gravedad son dos fuerzas que intervienen en la biosfera. No sólo conducen a una diferenciación vertical, sino que constituyen también la base de un mecanismo de transporte. En este sentido, el conjunto de la biosfera, que podría imaginarse como una película que cubriera el planeta, comparte muchas de las funciones de un modelo general de membrana biológica.

En los ecosistemas acuáticos, la sucesión conduce a la subdivisión de la columna vertical a través de múltiples discontinuidades. El proceso de diferenciación de una termoclina es un excelente ejemplo de estructuras que se autopotencian. En un lago, una situación algo complicada, pero usual, incluiría los siguientes compartimentos: atmósfera, epilimnion, hipolimnion, monimolimnion y sedimento. Esta estructura, a medida que se desarrolla, frena la tasa de recambio general y aumenta la segregación de las especies y la riqueza en detalle de la organización. Probablemente el desarrollo de esta organización frene el cambio de la razón entre el incremento de entropía y la cantidad de información o de organización acumulada para el futuro.

A lo largo del eje vertical tenemos un gradiente pronunciado, expresión de la energía disponible para crear y mantener una organización más bien rica, en forma de sistema disipativo. El plano horizontal es esencialmente distinto, y es apropiado modelar o concebir el ecosistema entero como si estuviera formado por una gavilla de columnas paralelas y comparables. En cualquier modelo que pretenda ser realista, la sección de cada columna vertical no puede ser infinitamente pequeña. La entrada de energía externa requiere una superficie mínima, relativamente pequeña en el caso de las olas, mucho mayor en los patrones marinos de la circulación asociada a las surgencias. Quizás ilustre comparar el ecosistema con una batería de tubos de ensayo dispuestos en paralelo en los que se esté llevando a cabo una reacción de Zhabotinskii; el conjunto puede llegar a marcar un ritmo. En el caso del ecosistema oceánico, el ciclo diario da lugar, al menos, a una onda en la intensidad de la producción primaria, que se refleja en la distribución de muchas propiedades en cada una de las columnas. Otra limitación en el tamaño de las columnas unidad del modelo depende de la presencia de animales capaces de migrar horizontalmente, generadores de un transporte y una diferenciación horizontales, entre otros factores.

Podemos asimilar cada columna vertical elemental de un ecosistema a una cuerda tensa y bien afinada. Al vibrar, se propagan las ondas a su alrededor y a través de cualquier plano horizontal. Los ritmos diarios, o anuales, del ecosistema generan ondas con distintas propiedades (químicas, bióticas, etcétera) que se desplazan a su alrededor. Las ondas internas de los ecosistemas acuáticos no se limitan a la temperatura y al movimiento del agua. La heterogeneidad horizontal que se desarrolla por interacción horizontal se muestra siempre menos clara que la organización vertical. Un ecosistema acuático equivale a un experimento de Zhabotinskii realizado sobre un plato. La falta relativa de determinación a lo largo del plano horizontal, a diferencia de la organización vertical, más rígida, explica los patrones fugitivos y aparentemente estocásticos que hacen tan difícil, y a la vez tan apasionante, el análisis de los transectos horizontales a través de las aguas de un lago o del océano. A menudo, la heterogeneidad que se observa a lo ancho de un plano horizontal permite reconocer un modelo en el que se distinguen dos partes topológicamente diferenciadas. Una parte son manchas discontinuas de agua más mezclada o turbulenta, poblada de organismos oportunistas que podrían considerarse representantes de los primeros estadios de una sucesión. El retículo, o panal, que rodea a estas manchas presenta, respecto al agua, un carácter más estable y, respecto a la sucesión, un carácter más avanzado, con estructuras más ricas y con movimientos más regulares de los organismos. Como suele ser habitual en la organización de la biosfera, las estructuras de distintos tamaños se combinan y solapan a diferentes escalas, desde la menor de las manchas al mayor de los bucles de la circulación marina.

Los cambios que se dan en los ecosistemas son de dos tipos. Uno discurre por medio de importantes entradas de energía procedente del exterior y conduce a distribuciones más bien uniformes, un resultado que podría predecirse a partir de la termodinámica elemental. En otros senderos de cambio, la energía también se degrada, pero, al mismo tiempo, la segregación de los reactivos, o de lo que quede de ellos, va en aumento, y el sistema se hace más heterogéneo. Los biólogos se fascinan ante esta modalidad de cambio, relativamente lento y en deceleración. Curiosamente han asociado este tipo de cambio con la idea de progreso. Quizá porque discurre en contra de ciertas tendencias atribuidas a los sistemas físicos. En cualquier caso, se encuadra en el área de estudio de la termodinámica

de sistemas abiertos lejos del equilibrio. La forma principal de estar fuera de equilibrio es dividir cualquier organización en individuos y poblaciones discontinuas, algo que la vida ha efectuado desde un principio. El estudio del plancton ofrece ejemplos muy instructivos y sugiere aproximaciones positivas. Muestra también que las dimensiones vertical y horizontal no son comparables en cuanto concierne a la organización de los ecosistemas. La organización vertical puede responder al stress por medio de un comportamiento oscilatorio. Si el ecosistema se descompone idealmente en numerosas columnas verticales adyacentes, el cambio horizontal se propaga de una columna a la otra en forma de ondas. La tendencia natural se orienta hacia la segregación local a lo largo del plano horizontal. Sin embargo, las entradas de energía pueden discurrir en contra de esa tendencia, pueden potenciar el transporte horizontal y dar lugar a sistemas altamente reactivos. (Ramón Margalef.)

Protolenguaje celular

Hubo un tiempo, no hace mucho de ello, en que las hormonas eran hormonas y los neurotransmisores, neurotransmisores. Una hormona era una sustancia secretada por un tipo específico de célula de una glándula endocrina, que se liberaba a la corriente sanguínea para que actuara sobre células receptoras específicas localizadas en algún punto del organismo de un vertebrado. A su vez, un neurotransmisor era un mensajero químico que transportaba las señales de una neurona a otra célula nerviosa o a una fibra muscular. Aunque se sabía que había neurotransmisores hasta en los animales pluricelulares simples, se consideraba las hormonas como un producto de la evolución de los vertebrados.

Los límites entre unas y otros se han difuminado. Ciertas hormonas no se sintetizan sólo en glándulas endocrinas, sino también en otros tejidos, incluso en muchas células cancerosas. Algunas se elaboran en células nerviosas y actúan de neurotransmisores; también hay neurotransmisores que proceden de glándulas y desempeñan funciones hormonales. Se ha descubierto que ambos, hormonas y neurotransmisores, median, como "factores tisulares", en las interacciones locales entre células. Se acaba de confirmar la presencia de hormonas y neurotransmisores en organismos unicelulares, y hasta en bacterias se han detectado hormonas.

Al amparo de la nueva situación se ha propuesto una nueva teoría que pone en relación hormonas, neurotransmisores y otros mensajeros intercelulares, y que remonta su ascendencia hasta las primitivas moléculas mensajeras de organismos unicelulares. La teoría y gran parte de los datos en que se fundamenta han nacido al socaire del laboratorio de Jesse Roth, del Instituto Nacional (norteamericano) de Artritis, Diabetes y Enfermedades Digestivas y Renales. Roth y sus colegas Derek LeRoith, Joseph Shiloach, James L. Rosenzweig, Maxine A. Lesniak y Jana Havrankova han resumido en *The New England Journal of Medicine* los datos, la teoría y la potencia de ésta para explicar el estado actual de nuestros conocimientos sobre comunicación intercelular.

Roth y sus colaboradores se han dedicado a la búsqueda de hormonas peptídicas en organismos unicelulares. (Se llama péptido a una breve cadena de aminoácidos, los sillares de las proteínas, que se sintetiza directamente a partir de la información genética.) La presencia de una hormona peptídica en un organismo indica que ese organismo ha heredado, y expresa, el ADN que codifica la hormona. (Hay otros tipos de hormonas, pero pueden fabricarse a partir de moléculas ya existentes en las células, por lo que su presencia pudiera ser accidental.) En el proceso de detección, los investigadores cultivan los organismos en un medio simple, los tratan con agentes diseñados especialmente para extraer un determinado péptido y detectan la presencia del péptido mediante anticuerpos radiactivos que se unen a él. Se purifica la hipotética hormona y se confirma su identidad ensayando su actividad biológica, esto es, comprobando si la hormona actúa como lo hace en un vertebrado.

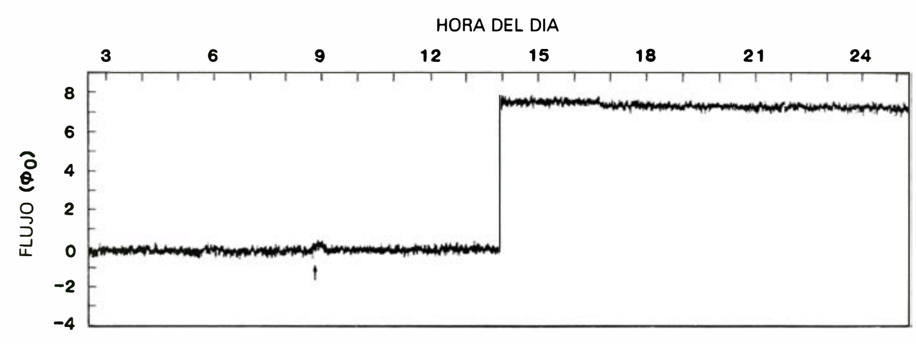
El grupo de Roth ha obtenido, en un protozoo, dos especies de hongos y tres cepas de la bacteria *Escherichia coli*, un producto indistinguible de la hormona peptídica insulina. En protozoos han hallado también somatostatina, péptido elaborado en el cerebro y páncreas de los mamíferos que inhibe la síntesis de insulina y diversas sustancias reguladoras, así como la hormona ACTH, corticotropina. Algunos investigadores habían informado con anterioridad de la presencia en protozoos de varios neurotransmisores propios de mamíferos, y de gonadotropina coriónica humana, otra hormona peptídica, en bacterias. Se desconoce por el momento cuál sea la función de esas moléculas mensajeras en organismos unicelulares, aunque hay indicios de que se unen al mismo tipo de moléculas receptoras que en los vertebrados.

Según Roth y colaboradores, los agentes bioquímicos de los sistemas nervioso y endocrino de los vertebrados, así como otros mensajeros intercelulares, se originarían en organismos unicelulares. Con el transcurso de la evolución, los elementos anatómicos (las células que producen mensajeros, los vasos y conductos que los transportan y las células que los reciben) se han ido diferenciando, se han especializado y ganado en complejidad. Sin embargo, la bioquímica de los agentes mensajeros se ha conservado en gran parte.

Afirman los investigadores que este enfoque de la historia de los sistemas mensajeros explicaría muchas observaciones recientes efectuadas en sistemas vertebrados. Da cuenta del solapamiento de las glándulas endocrinas y exocrinas, tanto en lo que se refiere a tipos celulares como a secreciones. (Las glándulas exocrinas descargan sus productos en una superficie, así los revestimientos estomacal e intestinal, y no en la corriente sanguínea.) Explica la presencia, en los vertebrados, de sustancias del tipo de los factores de crecimiento, prostaglandinas o interferón, que se asemejan a los péptidos endocrinos y cerebrales, si bien no se trata de hormonas ni de neurotransmisores. De acuerdo con esta hipótesis, todas las moléculas mensajeras fueron, en tiempos, factores tisulares locales, que actuaban sobre las células que los secretaban o sobre sus vecinas; un subconjunto de esas moléculas derivaría en hormonas y neurotransmisores. La teoría aclara incluso la situación de las feromonas, mensajeros químicos que libera un animal y afectan a la conducta de otro individuo de su misma especie. Las feromonas se equiparan en cierto grado a las hormonas; de hecho, cuando sólo había organismos unicelulares, cualquier mensajero intercelular era, por definición, una feromona.

¿Cayó, por fin, el monopolo?

A las 13,53 del 14 de febrero, día de San Valentín, se registraba un acontecimiento sin precedentes en un laboratorio de física de la Universidad de Stanford. El suceso, no del todo inesperado, hizo saltar un cepo minuciosamente preparado: el paso del acontecimiento quedó marcado en un drástico crecimiento de una corriente eléctrica persistente, puesta en movimiento meses antes en una bobina de hilo de niobio superconductor. El carrete estaba protegido de campos electromagnéticos externos por un cilindro de plomo superconductor. Todo el dispositivo se hallaba sumergido en un



Salto que delató el paso del monopolo magnético, según Blas Cabrera. Ocurría ello a las 13,53 del 14 de febrero de 1982

tanque de helio líquido mantenido a una temperatura de 4,2 grados Kelvin, por debajo de la temperatura a la que el plomo y el niobio pierden toda resistencia a una corriente eléctrica.

En las prudentes palabras de Blas Cabrera, el físico que preparó el cepo, el acontecimiento registrado "es concorde con el paso de una carga de Dirac solitaria". Dicho con términos menos esotéricos, la señal pudo producirla el paso de un monopolo magnético, la partícula de carga magnética aislada cuya existencia predijera hace ya más de medio siglo Paul A. M. Dirac.

En su breve informe de lo que aconteció ese día de San Valentín aparecido en *Physical Review Letters* razonaba los argumentos que le permitían deducir que ese salto de corriente debíase a la detección de un monopolo de rayo cósmico libre. Los esfuerzos dirigidos a la construcción de una gran teoría unificada de las fuerzas de la naturaleza incorporan la idea de que en los primeros 10^{-35} segundos subsiguientes a la gran explosión que originó el universo se crearon monopolos magnéticos superpesados. Ese tipo de monopolos portaría una masa excesiva y avanzaría con pesada lentitud, lo que impediría su registro en los detectores de partículas al uso, que descansan en las trazas de ionización dejadas por las partículas que corren casi a la velocidad de la luz. Ahí estaría la razón de por qué no se les había detectado hasta ahora, a pesar de los múltiples intentos por conseguirlo [véase "Monopolos magnéticos superpesados", por Richard A. Carrigan, Jr., y W. Peter Trower; INVESTIGACION Y CIENCIA, junio].

¿Qué método ha hecho triunfar a Cabrera? El basado en la interacción electromagnética de largo alcance de una carga magnética en movimiento con el estado cuántico macroscópico representado por el anillo superconductor. El detector empleado por el joven físico descendiente de españoles (su abuelo Blas fue un destacado especialista en magnetismo y su padre Nicolás lo es en estado sólido) incorporaba un sistema seguidor de corriente llamado SQUID (siglas de la expresión inglesa que designa al aparato de interferencia cuántica superconductora) y era insensible a la velocidad, dirección, masa y carga eléctrica de las partículas que lo atravesaran. Ni se inmutaría ante el paso de un dipolo magnético, formado por un polo norte y un polo sur. Pero el paso de una carga magnética solitaria a través del anillo provocaría un salto brusco, muy definido, de la corriente circulante.

Cuenta Cabrera que su detector ha estado funcionando, sin apenas solución de continuidad, durante más de 200 días. A lo largo de todo ese intervalo, sólo un acontecimiento, el del día de los enamorados, destacó lo suficiente como para que lo produjera el tránsito de un monopolo. El resultado pone una cota superior al número de partículas de carga magnética que corren por la superficie de la tierra. Límite que, a su vez, se acerca notablemente a lo calculado a partir de la hipótesis de que los monopolos magnéticos superpesados vendrían a contener toda la "masa perdida" de nuestra galaxia. Hipótesis que prevé, además, la detección de suceso y medio por año en un aparato del estilo del anillo de Cabrera. Pero lo que se impone es proseguir la búsqueda con este detector, confiados en atrapar otro monopolo en los próximos meses. Mientras se está en ello, puntualiza Cabrera, "hay en fase de construcción otros dos nuevos sistemas de una superficie sensible mayor".

Si se confirma el hallazgo de Cabrera con nuevas detecciones, las implicaciones que encierran son inmensas. Para W. Peter Trower, un empedernido cazador de los esquivos monopolos, del Instituto Politécnico de Virginia, "el experimento de Cabrera es elegante y admite muchas explicaciones distintas su episodio único. Aunque la prudencia nos dicta no aventurar ningún juicio, la verdad es que todos estamos muy ilusionados".

Efusiones de energía en estrellas jóvenes

La radiación emitida por las moléculas de monóxido de carbono en las nubes de gas donde se están formando estrellas revela masas de gas moviéndose hacia fuera a grandes velocidades

Charles J. Lada

En nuestra galaxia, las estrellas nacen en nubes de gas, cuerpos tan fríos que no emiten radiación alguna en las longitudes de onda visibles del espectro electromagnético. Las nubes están inmersas, además, en polvo cósmico, de modo que la radiación visible emitida por las nuevas estrellas queda absorbida. Estas condiciones opusieron durante largo tiempo una barrera, impenetrable por lo que parecía, a los astrónomos que trataban de comprender cómo se formaban las estrellas. Ahora bien, la radiación a las longitudes de onda del infrarrojo y las ondas de radio más cortas sí pueden atravesar esas nubes. Gracias al desarrollo de nuevos telescopios y equipo para la detección de radiación a esas longitudes de onda, los astrónomos disponen ya de los medios necesarios para explorar las nubes, oscuras y polvorientas, donde se gestan las estrellas.

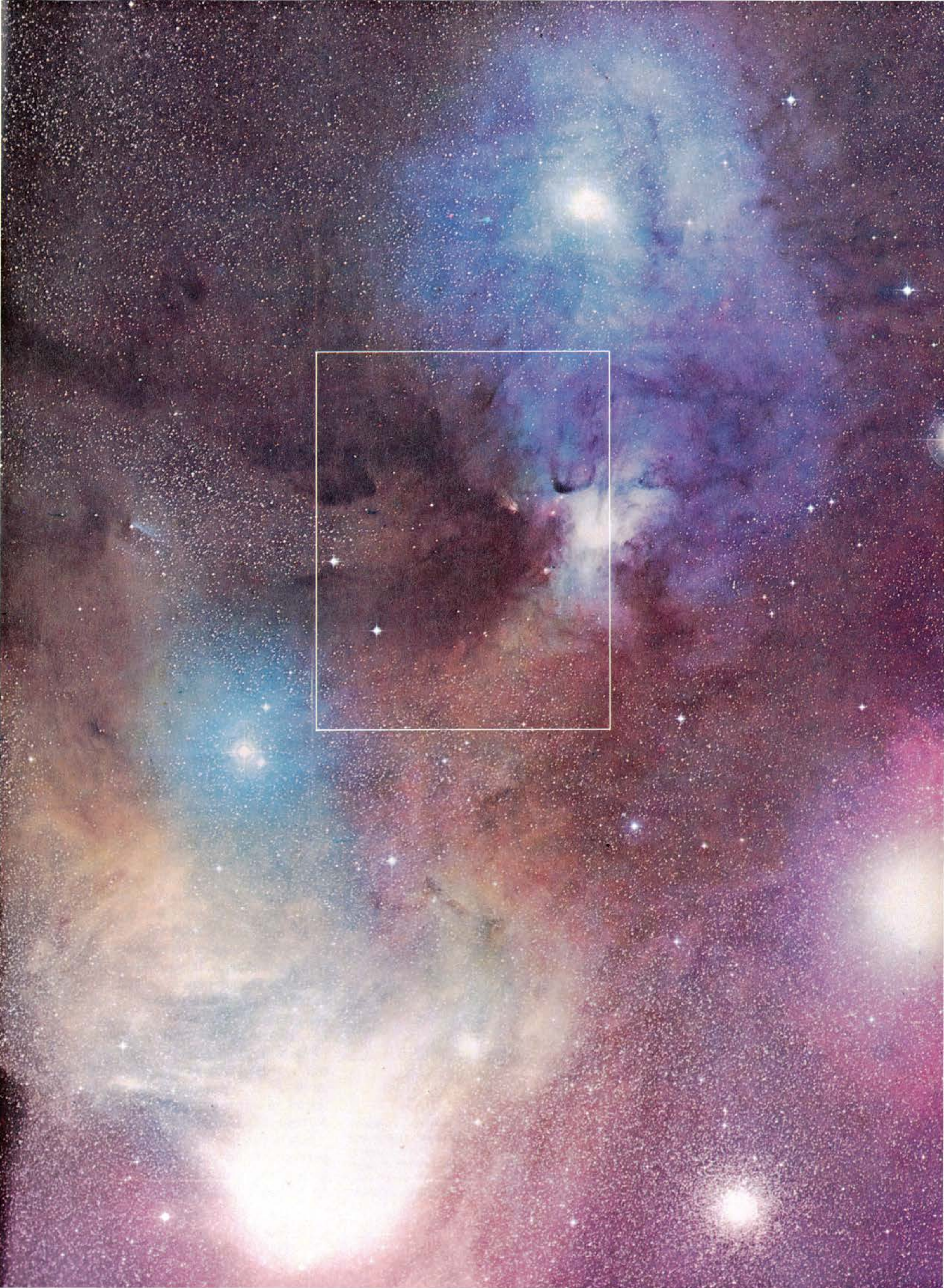
La molécula de monóxido de carbono (CO), gas detectado por primera vez en el espacio hace sólo una docena de años, ha demostrado ser una sonda excepcional para deducir las condiciones reinantes en el interior de nubes de donde saldrán las estrellas. En las nubes moleculares interestelares, las moléculas de monóxido de carbono emiten radiación a una longitud de onda de 2,6 milímetros. La investigación de tal radiación acaba de revelarnos un fenómeno astrofísico nuevo y fascinante, estrechamente relacionado con el nacimiento y evolución de las estrellas. Cuando ciertas estrellas están en las primeras fases de su vida, se las observa asociadas a violentas expulsiones de masa. En múltiples ejemplos, el gas molecular está saliendo de la vecindad de estrellas, de reciente creación, en dos chorros supersónicos que forman un ángulo de 180 grados. En esas corrientes bipolares, el gas tiene una masa de muchas veces la solar; lleva, en consecuencia, cantidades enormes de energía cinética. El origen y la naturaleza de esas efusiones de energía constituyen un misterio, por más que sepamos que representan una fase importante en la evolución de nuevas estrellas.

Los nuevos atisbos del nacimiento y evolución temprana de las estrellas se fundan en los datos suministrados por media docena de nuevos instrumentos especialmente proyectados para detectar la radiación en la región del infrarojo próximo del espectro (longitudes de onda de 2 a 20 micrometros), en el infrarrojo lejano (de 30 a 300 micrometros) y en las regiones de las ondas de radio más cortas (desde menos de un milímetro hasta algunos milímetros). Entre esos dispositivos citemos el telescopio infrarrojo de tres metros del Proyecto de Utilización, en Mauna Kea, Hawai, y el telescopio de cinco metros de onda milimétrica en Monte Locke, Texas. La superficie recubierta de oro del telescopio de cinco metros es el radiorreflector de onda milimétrica que más se aproxima a su forma ideal en todo el mundo. Hará tres años, el instrumento detectaba la primera corriente bipolar de alta velocidad en una nube molecular.

Pero, puestos a señalar a uno de entre esa nueva familia de instrumentos, citaríamos el Telescopio de Espejo Múltiple (MMT) del Monte Hopkins, cerca de Tucson, Arizona, de cuyo funcionamiento son responsables conjuntamente la Universidad de Arizona y el Observatorio Astrofísico Smithsoniano. Este instrumento único consta de seis espejos de 1,8 metros montados dentro de una circunferencia de siete metros. Cuando opera de suerte tal que la radiación incidente sobre los seis espejos se combina de manera coherente, es decir, en fase, se convierte en el equivalente de un reflector de siete metros. Recientemente, se hizo que el MMT funcionara en el modo coherente para examinar las emisiones de nubes moleculares a longitudes de onda de menos de un milímetro, región del espectro en la que, hasta ahora, había sido difícil realizar observaciones de alta resolución.

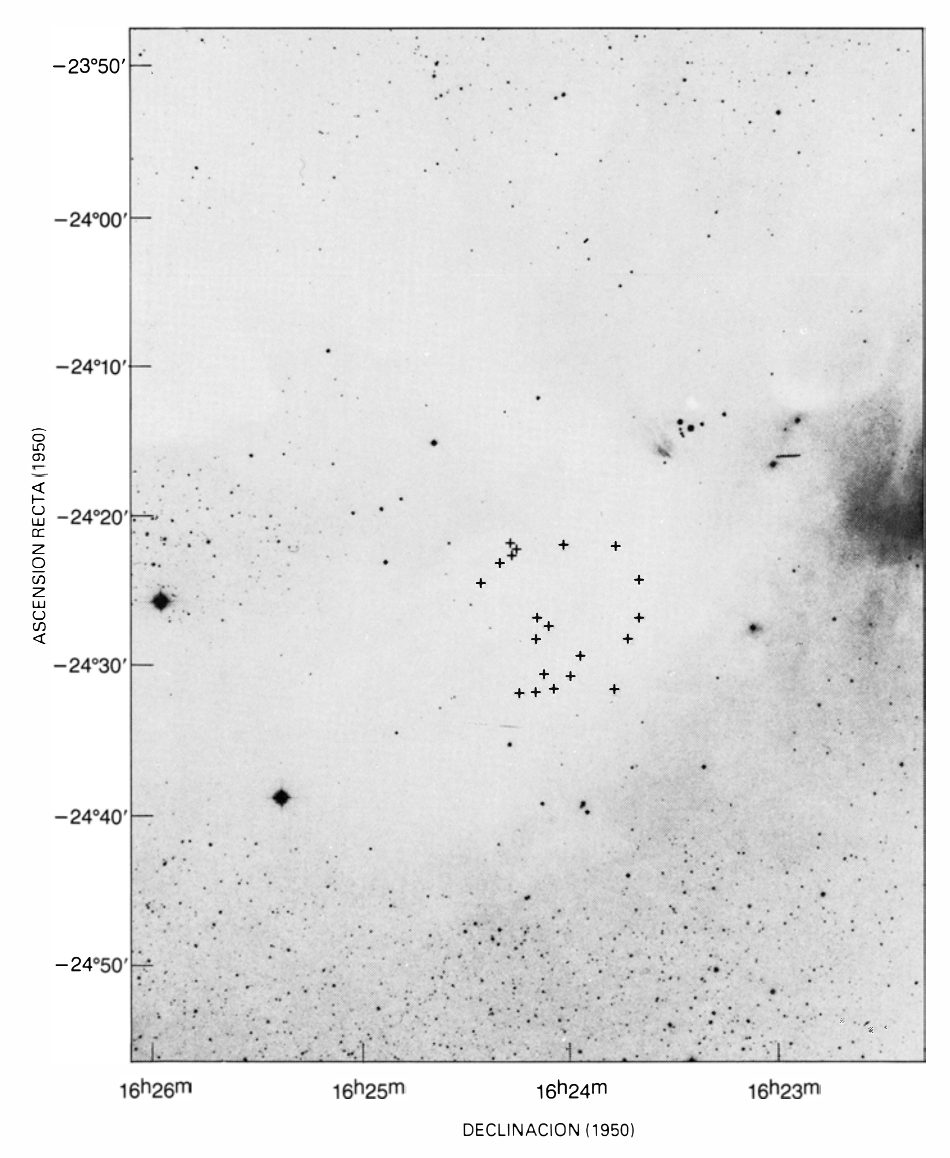
Los estudios de las emisiones de líneas específicas en la región del espectro de ondas milimétricas han posibilitado que los astrónomos sondeen la naturaleza física, química y dinámica de las nubes oscuras a partir de las cuales se están formando estrellas. A las observaciones en el infrarrojo lejano revelan el estado físico del polvo que impregna las nubes oscuras y oculta las estrellas recién nacidas. A las observaciones en el infrarrojo próximo debemos también la revelación de estas estrellas. Así, en la nube oscura co-

UN LUGAR TIPICO DE NACIMIENTO DE ESTRELLAS es la oscura nube Rho de Ofiuco, llamada así por la brillante estrella rodeada de una nebulosidad de luz azul reflejada en la parte superior central de esta fotografía. La brillante estrella de la parte inferior izquierda es Antares, en la constelación del Escorpión. A la derecha de Antares se encuentra el cúmulo globular M4. La nube oscura entre Rho de Ofiuco y Antares contiene hidrógeno molecular gaseoso cuya masa equivale aproximadamente a la de 100 veces el Sol. Densas concentraciones de polvo en el interior de la nube molecular absorben más del 99,99 por ciento de la radiación visible emitida por un cúmulo de 20 nuevas estrellas, descubierto recientemente en una exploración hecha en longitudes de onda del infrarrojo por Bruce A. Wilking, de la Universidad de Texas en Austin, y el autor. La radiación infrarroja multiplica por diez la eficacia de la visible en cuanto a su penetración a través del polvo. El cúmulo de nuevas estrellas está oculto cerca del centro del área cerrada en recuadro. En la página 46 hay una fotografía de ésta que señala la posición de los objetos infrarrojos. La fotografía fue realizada por David F. Malin a partir de tres tomas en blanco y negro expuestas a distintas regiones del espectro con el telescopio Schmidt de 1,2 metros en Siding Spring.



nocida por Rho de Ofiuco, las observaciones en el infrarrojo han descubierto todo un cúmulo de 20 estrellas, de reciente creación, profundamente inmersas en gas y polvo, hasta el punto de que nunca se hubieran detectado en longitudes de onda visibles. La síntesis de datos procedentes de esas diversas regiones del espectro correspondientes a longitudes de onda invisibles ha prestado un favor impagable al conocimiento que podamos tener ahora del nacimiento y primeras fases de la evolución estelar. A su debido tiempo, tal trabajo resultará, como fruta madura, en una teoría general de la formación de estrellas.

No hay que remontarse muy atrás para recordar que el tema de la formación estelar y el de la génesis de sistemas planetarios se movían en el terreno de la pura especulación teórica. No existían datos de observación pertinentes. Hace 40 años, la situación cambió de golpe merced a los avances registrados en física nuclear y dinámica estelar. Quedó claro que la provisión de energía de las estrellas venía de la fusión nuclear. Reconocido lo cual, cabía predecir cuánto tiempo podrían vivir las estrellas y cual debería ser su destino final. Viose así que las estrellas muy luminosas designadas *O* y *B* estaban consumiendo su reserva de combustible nuclear de hidrógeno de 10.000 a 100.000 veces más deprisa que el Sol; valía decir que estaban condenadas a agotar su combustible nuclear en unos 10 millones de años o antes. Tales estrellas no llegarán a vivir ni siquiera una rotación de nuestra galaxia (100 millones de años) y tienen una vida previsible de una milésima parte escasa de la solar. Las estrellas *O* y *B*, muy jóvenes, no pueden haberse alejado mucho de sus lugares de nacimiento en las nubes de gas y polvo.



NUEVAS ESTRELLAS EN LA NUBE OSCURA Rho de Ofiuco identificadas por cruces en un negativo del área enmarcada en blanco en la página anterior. B. A. Wilking y el autor exploraron las regiones más oscuras de la nube a una longitud de onda de 2 micrometros con el telescopio infrarrojo de 1.5 metros de la Universidad de Arizona. Se estima que se han formado 20 estrellas en el último millón de años. La fotografía, tomada con el telescopio Schmidt de 1.2 metros de Monte Palomar, procede de la Exploración del Cielo conjunta de la National Geographical Society y el Observatorio de Monte Palomar.

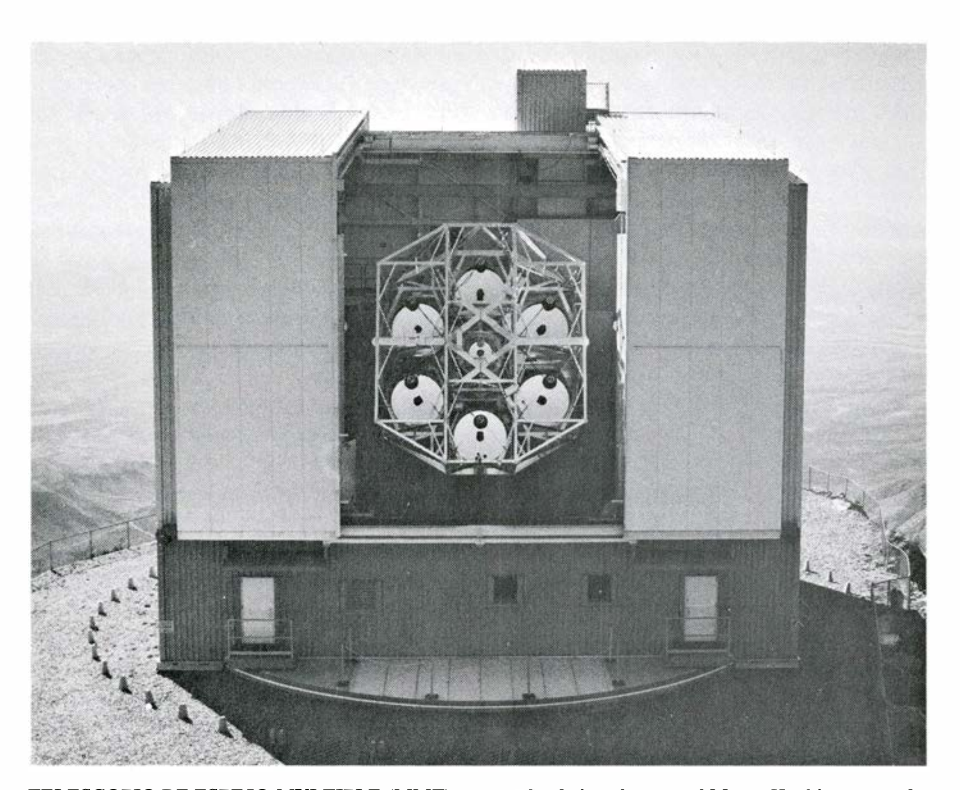
Lugares de nacimiento de estrellas

Los estudios de la distribución y movimiento de las estrellas *O* y *B* han puesto de manifiesto que muchas se integran en grupos, las llamadas asociaciones *OB*. No se ligan gravitatoriamente, como ocurre entre las estrellas de un cúmulo, sino que los miembros de una asociación *OB* se alejan, en su movimiento, unas de otras. La velocidad de expansión indica que las asociaciones no tienen una antigüedad superior a los 10 millones de años. Estas dos piezas independientes de información —la alta velocidad de consumo de combustible de las estrellas *O* y *B* y su movimiento de alejamiento— muestran que, todavía hoy, brotan estrellas en el seno de nuestra galaxia. Por ser tan breve el ciclo de vida de las *O B*, desde su nacimiento hasta su muerte, su distribución y orientación en una pequeña región del espacio se puede considerar como un archivo fósil de la distribución y orientación de las protoestrellas que fueron sus precursoras y presumiblemente del gas y polvo que precedió a las protoestrellas.

La llegada de las observaciones en las regiones infrarroja y milimétrica del espectro ha revelado que las asociaciones *OB* surgen en nubes moleculares gigantes de unos 300 años-luz de diámetro, con una masa de unas 100.000 veces la solar, formando un gas más tenue que el mejor vacío que pueda obtenerse en el laboratorio. Quizá sean tales agregaciones gaseosas los objetos de mayor masa de la galaxia. Hace doce años, su existencia era desconocida [véase «Complejos gigantes de nubes moleculares en la galaxia», por Leo Blitz, INVESTIGACION Y CIENCIA, junio]. Se acepta ya, fundados sobre todo en las observaciones en las regiones infrarroja y milimétrica, la posibilidad de varios mecanismos para la formación de estrellas a partir de las nubes moleculares gigantes. ¿Cuántos, exactamente? ¿Cuál es su contribución relativa a la formación de estrellas? Esas cuestiones, que polarizan la atención de gran parte de la investigación astronómica actual, ni siquiera podrían haberse formulado diez años atrás.

Las nubes que originan las estrellas parecen constar casi exclusivamente de hidrógeno molecular (H_2). Por reinar en las primeras una temperatura baja

(entre 10 y 50 grados Kelvin, o grados Celsius por encima del cero absoluto), no se pueden detectar directamente las moléculas de hidrógeno. En el espacio sólo se detectan las moléculas que emiten fotones, lo que acontece, por ejemplo, al caer de un estado energético a otro más bajo. Las transiciones de un estado a otro dan lugar a líneas espectrales, que, dependiendo de la energía perdida al efectuar la transición, pueden caer en cualquier parte del espectro electromagnético. En los gases moleculares a baja temperatura, las principales transiciones se presentan en un estado de rotación de la molécula a otro estado, transiciones en que la radiación se emite a radiofrecuencia. De acuerdo con las reglas de la mecánica cuántica, para una molécula simétrica, como la de hidrógeno, están prohibidas las transiciones entre niveles de energía rotacional vecinos. La primera transición por encima del estado fundamental (o más bajo) permitida para el hidrógeno molecular requiere, pues, que la molécula gane energía suficiente para subir dos niveles de rotación. A las temperaturas predominantes en las nubes moleculares, los choques entre moléculas raramente suministrarán la energía requerida para tal transición.



TELESCOPIO DE ESPEJO MULTIPLE (MMT), que acaba de instalarse en el Monte Hopkins, cerca de Tucson, Arizona. Es el primero de su clase. En su modo normal de funcionamiento, conduce hasta un foco común la radiación procedente de seis espejos de 1,8 metros montados en una circunferencia de siete metros de diámetro, con lo que iguala el poder captador de luz de seis telescopios de 1,8 metros o un telescopio único de 4,4 metros. Sin embargo, como la luz se combina en forma incoherente, el poder de resolución del telescopio de espejo múltiple sigue siendo el de un instrumento de 1,8 metros. Para alcanzar el máximo poder de resolución del MMT, su luz debe combinarse en forma coherente, o en fase. Cuando se hace esto, el MMT equivale a un telescopio de siete metros de diámetro. Se empleó por primera vez un sistema de fase coherente a longitudes de onda submilimétricas usando el MMT por Neal Erickson, Paul F. Goldsmith, C. Richard Huguenin y Ronald L. Snell, de la Universidad de Massachusetts en Amherst, en colaboración con Bobby L. Ulich, del observatorio del MMT, y el autor del presente artículo.

Detección del gas molecular

Por fortuna, las nubes moleculares pueden también contener moléculas no simétricas, monóxido de carbono por ejemplo; en cuyo caso sí están permitidas las transiciones entre niveles de rotación vecinos. En las propias nubes más frías, los choques entre moléculas elevan sin dificultad las moléculas de monóxido de carbono hasta su primer estado de rotación excitado. Cuando la molécula cae de nuevo a su estado fundamental, emite fotones de una longitud de onda de 2,6 milímetros. Aunque la abundancia de monóxido de carbono en su forma isotópica común (carbono 12 combinado con oxígeno 16) constituye la diezmilésima parte de la cantidad de hidrógeno molecular, es una de las principales moléculas de las nubes interestelares; sirve, por tanto, de excelente sustituto para abordar las inobservables moléculas de hidrógeno. La primera línea espectral de 2,6 milímetros del monóxido de carbono detectada en el espacio interestelar lo fue en 1970, por Robert Wilson, Keith B. Jefferts y Arno A. Penzias, de los Laboratorios Bell, con el radiotelescopio de 11 metros del Observatorio Nacional de Radioastronomía de Kitt Peak, en Arizona. Desde entonces se ha observado la línea espectral del monóxido de carbono en todos los puntos de nuestra galaxia, y en otras.

A partir de la intensidad de las líneas espectrales moleculares, los astrónomos deducen la temperatura, densidad y composición molecular de las nubes donde se forman las estrellas. Además, como consecuencia del efecto Doppler (que eleva la frecuencia de las fuentes radiantes que se están acercando al observador y reduce la frecuencia de las que se están alejando), llegan a determinar el movimiento de una nube molecular con respecto al sistema solar. De las fuentes que se acercan se dice que están corridas hacia el azul; de aquellas que se alejan, que lo están hacia el rojo. El efecto Doppler también pone de manifiesto el dominio de velocidades de las moléculas en una nube molecular; cuanto mayor el dominio de velocidades, tanto mayor la anchura de la línea espectral. La observación del perfil de líneas moleculares corridas y ensanchadas condujo al descubrimiento de unas corrientes gaseosas poco frecuentes en las nubes moleculares.

Si los movimientos térmicos fuesen la única fuente del ensanchamiento Doppler de las líneas, las nubes moleculares interestelares emitirían líneas con una anchura equivalente a sólo unas décimas de kilómetro por segundo. Las anchuras de las líneas suelen expresarse en unidades de velocidad y no de longitud de onda, ya que el máximo interés se centra en los movimientos de las nubes moleculares y de sus constituyentes. Las líneas de monóxido de carbono de las nubes moleculares son casi siempre mucho más anchas que las líneas que producirían, por sí solos, los movimientos térmicos.

Gas de alta velocidad en Orión

En el caso de las nubes moleculares gigantes, pensemos en las que se encuentran en la vecindad de las asociaciones *OB*, se observan, en la mayor parte de la nube, anchuras que corresponden al dominio de uno a tres kilómetros por segundo. En las regiones donde la formación de estrellas presenta su máxima intensidad, las anchuras aumentan hasta cuatro o 10 kilómetros por segundo. La fuente de tales movimiento está envuelta en el misterio.

El ejemplo que constituye el caso

más extremo de ensanchamiento de una línea molecular se descubrió hace unos cinco años en una pequeña región de una nube molecular gigante de 200 a 300 años-luz de diámetro, íntimamente asociada con la bella nebulosa de Orión. En una región de diámetro inferior a medio año-luz, la emisión de monóxido de carbono muestra un intervalo de velocidades ligeramente superior a los 100 kilómetros por segundo. Los estudios de esta pequeña región a muchas longitudes de onda muestran que la alta velocidad de sus moléculas de monóxido de carbono tiene que ver con un chorro de gas de gran energía, procedente de los alrededores de un objeto joven, una protoestrella tal vez.

La existencia de una tal efusión de materia gaseosa, de gran energía y posiblemente violenta, desde la vecindad de un objeto joven, no encajaba en el cuadro de la formación y evolución temperana de las estrellas hasta entonces aceptado. Se creyó al principio que el chorro guardaba relación con un suceso muy raro y único quizá: la explosión de una supernova en la nube molecular de Orión solamente unos 1000 años antes. Desde entonces, las observaciones en onda milimétrica de otras nubes moleculares han puesto en claro que tales efusiones de gas de alta energía no constituyen ninguna rareza y pueden ser incluso consecuencia del nacimiento de las estrellas.

La nube alrededor de AFGL 490

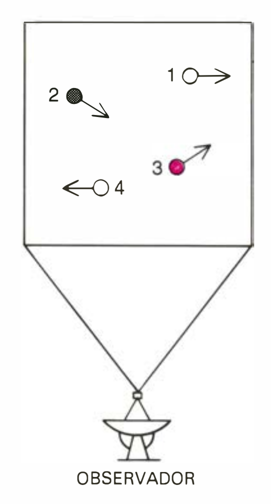
Paul M. Harvey, de la Universidad de Texas en Austin, y el autor realizaron, hace un par de años, una serie de observaciones de esas. El objeto de nuestro estudio era una nube molecular alrededor de una fuente infrarroja, de la que se sospechaba su carácter protoestelar, en la constelación de la Jirafa. Se la venía conociendo por su designación en dos catálogos infrarrojos: U de A1 (objeto n.º 1 en el catálogo de la Universidad de Arizona) y AFGL 490 (objeto n.º 490 en el catálogo del Laboratorio Geofísico de la Fuerza Aérea (AFGL)).

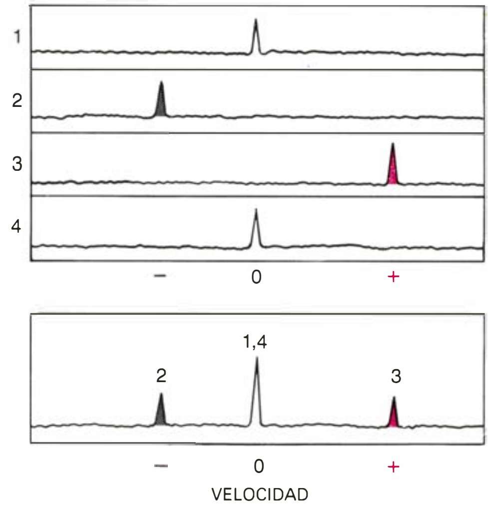
Acometimos nuestros primeros trabajos con el reflector de cinco metros, recubierto de oro, del Observatorio McDonald de la Universidad de Texas en Austin, sito en el Monte Locke. Cuando apuntamos el telescopio directamente hacia la fuente infrarroja oculta, nos sorprendió que el perfil del monóxido de carbono tuviera anchas «colas», debidas a velocidades. En vez de presentar la anchura esperada, de dos o tres kilómetros por segundo, se detectaba la emisión del monóxido de carbono a lo largo de un dominio de unos 60 kilómetros por segundo. En puntos de la nube distintos de la posición de la fuente infrarroja, las anchas colas debidas a la velocidad desaparecían, y las líneas ofrecían las pequeñas anchuras que en un comienzo nosotros esperábamos encontrar.

Tan sorprendente descubrimiento mereció ulterior investigación con un telecospio mayor y más sensible. Trasladamos nuestro trabajo al instrumento de 11 metros de Kitt Peak, que goza de una resolución doble de la que posee el telescopio de cinco metros. Mejor resolución que nos abrió la puerta a otro inesperado descubrimiento. Al barrer con el telescopio la fuente infrarroja de Norte a Sur, la silueta del perfil del monóxido de carbono variaba de una manera sistemática. Ligeramente al Norte de la fuente observamos sólo la cola corrida hacia el rojo del perfil. Ligeramente al Sur, sólo se percibía la cola corrida hacia el azul. Al centrar el telescopio directamente sobre la fuente infrarroja, el espectro del monóxido de carbono mostraba componentes de alta velocidad corridas tanto hacia el rojo como hacia el azul, cosa muy distinta de lo que se pensaba encontrar en una nube gaseosa que estuviera colapsando gravitatoriamente para formar una nueva estrella.

El chorro de gas se extiende a lo largo de un año-luz, o 0,3 parsec, en cualquier dirección. Debe tratarse, pues, de un suceso dinámico jovencísimo. (Un parsec equivale a 3,26 años-luz). El gas más veloz de esta corriente tardaría 10.000 años escasos en recorrer un año-luz. Las propias nubes moleculares tienen una vida estimada de 10 a 100 millones de años. Observando la emisión de una forma isotópica rara del monóxido de carbono (en la que el carbono era carbono 13, no 12), determinamos que la masa de hidrógeno en la corriente era de alrededor de 30 veces la masa solar. Un sencillo cálculo (la mitad de la masa multiplicada por el cuadrado de la velocidad) muestra que el gas está transportando alrededor de 2×10^{47} erg de energía cinética, equivalente a la emisión total de energía radiante del Sol durante 1,6 millones de años. Evidentemente, en la vecindad de AFGL 490 se acaba de producir un suceso de alta energía. ¿Qué razón dar de la actividad observada?

En principio, las corrientes movidas hacia el azul y el rojo podrían representar los bordes que se acercan y se alejan de un disco de gas en rotación que se dejara ver de canto. Para mantenerse estable frente a la tendencia centrífuga a destruirlo, el disco habría de estar constituido por al menos 20.000 masas solares de materia dentro del diámetro de 0,6 parsec de la emisión de alta velocidad. Nuestras observaciones señalaban que, dentro de esta región, la cantidad total de materia (incluyendo el gas de la corriente de alta velocidad y el gas ambiente de baja velocidad) tenía un límite superior de 100 masas solares. Además, las observaciones en el infrarojo del objeto central oculto, esto es, el propio AFGL 490, sugerían que su masa no superaba 15 veces la solar. La explicación más plausible de las corrientes de alta velocidad representadas por los corrimientos hacia el azul y ha-





MOVIMIENTO RELATIVO de las moléculas en el espacio. En virtud del efecto Doppler, dicho movimiento altera las longitudes de onda a las cuales se observa su emisión. Si la molécula está en resposo o se está moviendo transversalmente (*1, 4*), el "pèrfil" de su línea no se altera. Si se está acercando al observador (*2*), el "perfil" de su línea se trasladará hacia una longitud de onda más "azul", o más corta. Si se está alejando (*3*), su perfil se correrá hacia una longitud de onda "más roja". El corrimiento revela la velocidad. La intensidad de una velocidad Doppler depende de las moléculas que emiten a esa velocidad.

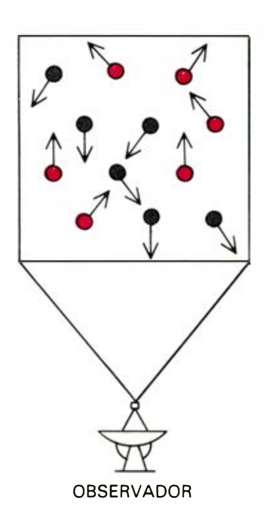
cia el rojo es, pues, que no se hallan confinadas gravitatoriamente. Antes bien, el objeto central las impulsa hacia fuera, por algún mecanismo, en dos chorros: uno del que podríamos decir que se dirige hacia nosotros y el otro que se aleja en sentido opuesto. Tal chorro bipolar no requería una gran concentración de materia en el objeto central a su alrededor.

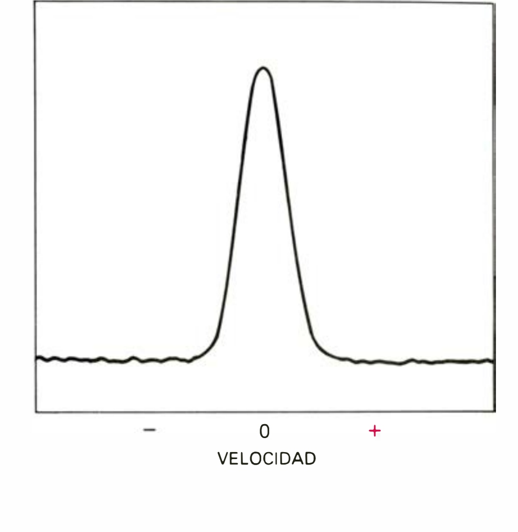
Los fenómenos de un chorro de gas de alta energía observados en la nube molecular de Orión y el AFGL 490 se han observado desde entonces en otros lugares. En el momento de escribir este artículo se han identificado 24 ejemplos más. De ellos se han cartografiado unos 12; casi todos presentan la estructura bipolar.

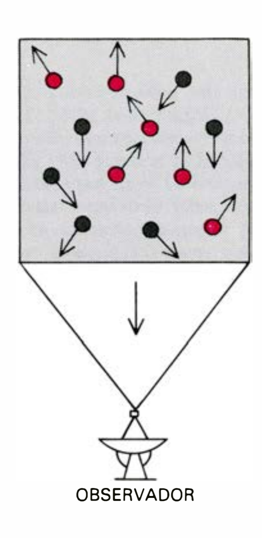
Los chorros moleculares de L1551

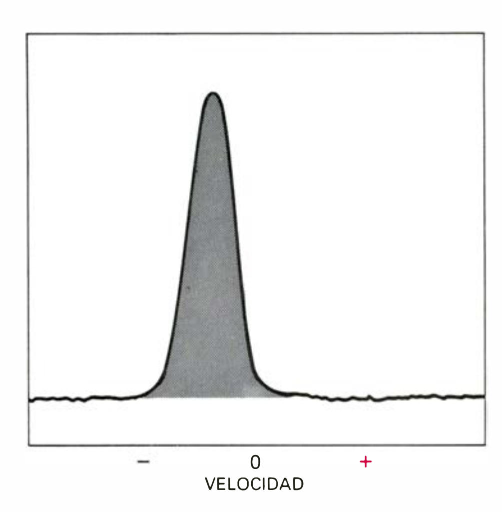
Tal vez el ejemplo más notable de chorro molecular bipolar sea el asociado a una fuente infrarroja en la oscura nube molecular L1551, de la constelación de Tauro. Ya las primeras observaciones sugirieron algo peculiar en los perfiles del monóxido de carbono de la vecindad de la fuente. Con el telescopio de cinco metros de Monte Locke, Ronald L. Snell, de la Universidad de Massachusetts en Amherst, Robert B. Loren, de la Universidad de Texas en Austin, y Richard L. Plambeck, de la Universidad de California en Berkeley, examinaron la fuente con cierto detalle. Descubrieron chorros de gas molecular que soplaban desde una estrella joven en configuración bipolar.

Morfología que resaltaba por dos aspectos, nada usuales. En primer lugar, se trataba del único sistema de corrientes en una nube molecular que se ha resuelto completamente por medio de observaciones. Las longitudes de las corrientes de alta velocidad son mayores que sus anchuras en la proporción aproximada de tres a uno, lo que sugiere unas corrientes estrechas y en forma de chorro. En segundo lugar, el chorro que está corrido hacia el azul coincide con tres manchas visibles de emisión nebulosa de las que se cree sean objetos de Herbig-Haro. Tales objetos (llamados así por sus descubridores, George H. Herbig, de la Universidad de California en Santa Cruz, y Guillermo Haro, del Instituto Nacional de Astrofísica, Optica y Electrónica de México) parecen haber sido creados por ondas de choque, que se producen cuando fuertes vientos estelares dirigidos hacia fuera de un objeto interaccionan con la nube de gas molecular circundante.





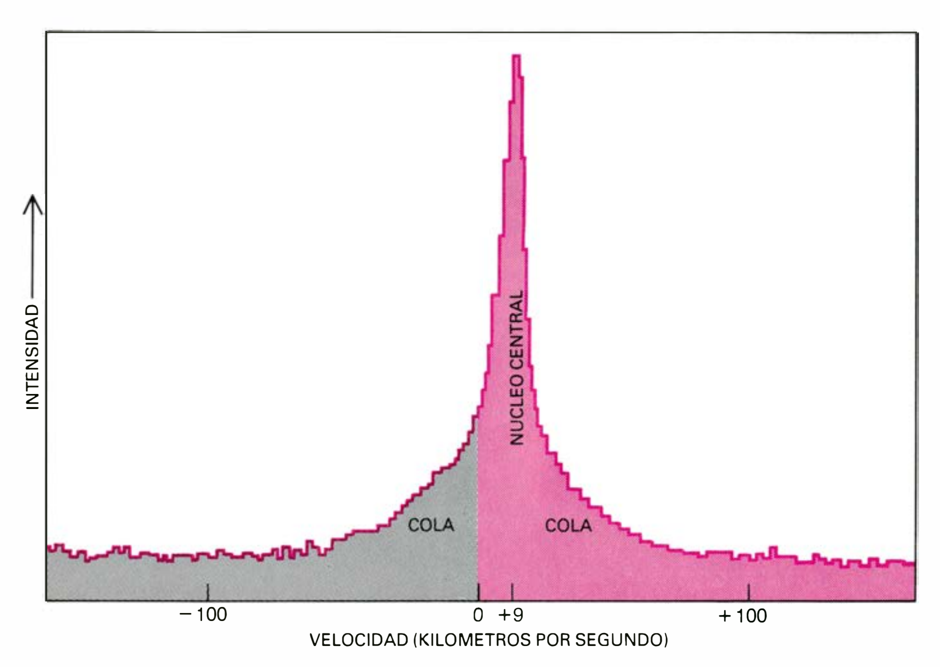




EN LA OBSERVACION REAL DE UNA NUBE MOLECULAR se suman las emisiones de las moléculas que se están moviendo dentro de cierto intervalo de velocidades. Si la nube no se está acercando ni alejando (*arriba*), la dispersión de velocidades debida a los movimientos internos conducirá a una línea de perfil ancho, centrada en la posición de la longitud de onda normal, o en reposo. Si la nube se está acercando (*caso inferior*), el perfil aparece corrido hacia el azul, lo que se describe como velocidad negativa. Una nube que se estuviera alejando mostraría un perfil corrido hacia el rojo, o velocidad positiva.

Las observaciones ópticas de las líneas de emisión de los objetos Herbig-Haro en L1551 realizadas por Stephen E. Strom y Karen M. Strom, del Observatorio Nacional de Kitt Peak, y Gary L. Grasdalen, de la Universidad de Wyoming, han mostrado que las líneas presentan un corrimiento hacia el azul concordante con el corrimiento hacia el azul del chorro molecular, donde los objetos parecen estar inmersos. Los objetos Hebig-Haro se están acercando, según se han encontrado, a lo largo de la visual a una velocidad 50 kilómetros por segundo mayor que la del gas en el chorro molecular. En observaciones separadas, W. J. Luyten, de la Universidad de Minnesota, ha distinguido dos objetos de Herbig-Haro, H-H 28 y H-H 29, que se mueven en una dirección transversal a la visual, y lo hacen con una rapidez tal que ocupan diferentes posiciones en placas fotográficas tomadas con sólo unos años de intervalo. Cuando Kyle M. Cudworth, del Observatorio de Yerkes de la Universidad de Chicago, y Herbig analizaron placas de la región L1551 tomadas a lo largo de un período de 30 años, hallaron que H-H 28 y H-H 29 tenían velocidades transversas de unos 150 kilómetros por segundo. Probaron asimismo que, si los movimientos de los dos objetos se extrapolaban desde entonces hacia atrás en el tiempo, sus trayectorias se cortarían 3000 años antes en un punto próximo al origen del chorro bipolar saliente de gas molecular.

La razón presumible de que los objetos de Herbig-Haro se contemplen asociados al chorro corrido hacia el azul, y no al corrido hacia el rojo, reside en que el gas corrido hacia el azul se halla más cerca de la superficie próxima de la



LA NUBE MOLECULAR DE ORION, que recubre un cúmulo de estrellas recientes, mostraba este perfil de emisión de monóxido de carbono cuando fue observada por autor y sus colegas con el MMT. La radiación, que tiene una longitud de onda en reposo de 0,87 milímetros, se emite cuando las moléculas de monóxido de carbono descienden desde su tercer estado excitado de rotación hasta el segundo estado excitado de rotación. La nube entera se está alejando a nueve kilómetros por segundo. Las anchas "colas" del perfil indican, sin embargo, que muchas moléculas del interior de la nube se están moviendo a velocidades supersónicas de más de 50 kilómetros por segundo, acercándose unas y alejándose otras. La emisión de alta velocidad, confinada a una región de sólo una décima de año-luz de extensión, representa una corriente de gas de alta energía que procede de jóvenes estrellas que se hallan inmersas en la nube.

nube molecular, generando una onda de choque que da lugar a las nebulosidades luminosas de Herbig-Haro cuando el gas comienza a fracturar la superficie. La corriente corrida hacia el rojo está dirigida en sentido opuesto a nosotros; se encuentra sumergida en la nube a mayor hondura; así pues, cualesquiera objetos de Herbig-Haro que puedan crearse en el chorro saliente quedan completamente ocultados por el gas molecular y el polvo interpuestos.

Actividad energética

Si hubiera de ponderar algún aspecto de la violenta actividad observada en L1551, destacaría la naturaleza del objeto presumiblemente responsable de ella. En 1975, Stephen Strom y Karen Strom, que trabajaban en equipo con Frederick V. Vrba, del Observatorio Lovell de Flagstaff, Arizona, descubrieron la primera fuente infrarroja cerca del centro del gas saliente. No obstante, la fuente difería notablemente de las restantes fuentes infrarrojas halladas antes en la nube molecular de Orión y en AFGL 490. Las fuentes que impulsan los chorros de gas en esos dos lugares parecen ser estrellas jóvenes o protoestrellas, quizás estrellas del tipo espectral *B*, que tienen una masa de 15 veces la del Sol y son al menos 1000 veces más luminosas. Habida cuenta de su masa y luminosidad, a nadie sorprenderá que se aprecie una actividad energética en sus alrededores.

Contrasta lo anterior con las observaciones en el infrarrojo próximo y lejano de la fuente en L 1551 llevadas a cabo por Charles Beichman, del Instituto de Tecnología de California, y mediante un telescopio transportado en globo por un grupo de astrónomos europeos (dirigidos por C. V. Fridlund, del Observatorio de Estocolmo), que muestran que la luminosidad de la fuente sólo multiplica por 25 el brillo solar. El objeto central se parece a una estrella del tipo T de Tauro, joven cuerpo estelar cuya masa quizá duplique la del Sol. ¿Qué conclusión sacar entonces? Que las estrellas extraordinariamente masivas y luminosas no son los únicos objetos capaces de suministrar la energía que necesitan los intensos chorros bipolares de gas en las fases tempranas de su evolución. Puede haber estrellas no mayores que el Sol capaces de inducir los mismos fenómenos al principio de sus vidas. De los 26 chorros de alta velocidad conocidos, cuatro se hallan verosímilmente asociados a estrellas de masa bastante baja.

La actividad energética desplegada por la nube molecular de Orión, AFGL 490 y L1551 representa un nuevo y extraordinario fenómeno estelar definido por tres características distintivas, y, con frecuencia, cuatro. La primera, un chorro de gas molecular de alta velocidad. En segundo término, la actividad dinámica presenta una escala de tiempo de evolución corta: 1000 y 10.000 años. Lleva consigo, en tercer lugar, grandes cantidades de energía cinética: de 10^{45} a 10^{48} erg. Y por último, en la mayoría de los casos, el gas saliente muestra una configuración bipolar centrada en una fuente de infrarrojo oculta que es una estrella joven.

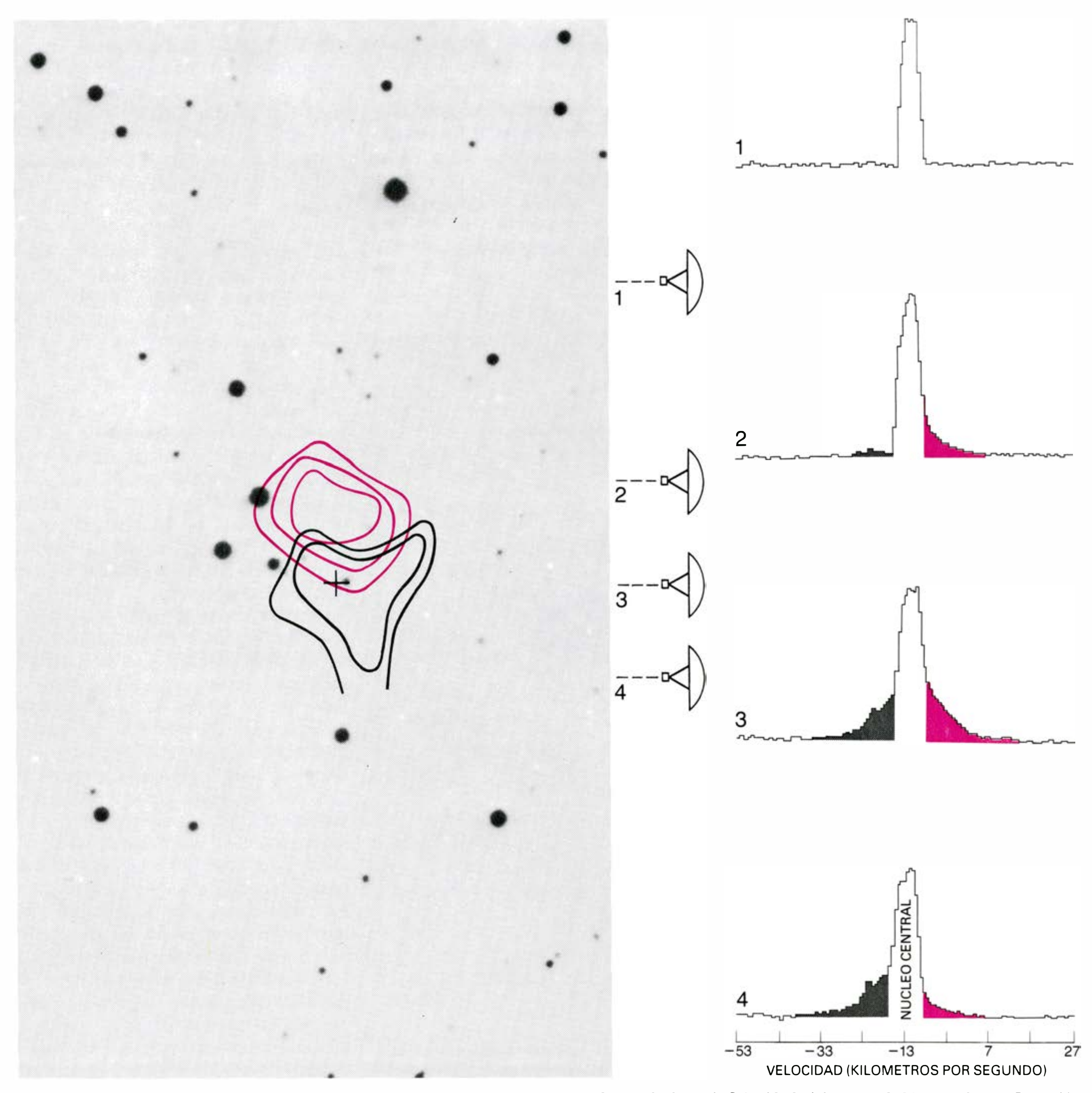
Frecuencia de los chorros moleculares

Dentro de un kiloparsec del Sol (3260 años-luz) se conocen ahora 10 ejemplos bien estudiados de este fenómeno. Los objetos se crean a velocidad notablemente alta: cada 10.000 años habrán de formarse 10 de ellos, por lo menos, sólo en la proximidad del sistema solar. La velocidad de formación de tales chorros moleculares centrados en una estrella viene a ser igual, según parece, a la velocidad de formación de estrellas con masas idénticas o superiores a tres veces la solar. Si consideramos que no se ha emprendido todavía ninguna búsqueda sistemática de chorros moleculares, la cifra asignada a su velocidad de formación será, probablemente, un límite inferior. A modo de conclusión provisional se avanza que todas las estrellas de algunas veces la masa solar atraviesan una fase temprana de su evolución en la que originan chorros de gas de gran masa y energía. Ignoramos si las estrellas de masa menor, parecidas al Sol, pasan por esa fase.

La posible existencia de semejante fase de evolución para las estrellas del tipo solar arrojará luz sobre el proceso de formación de planetas, del que resta mucho por conocer. Los planetas jovianos del sistema solar (Júpiter, Saturno, Urano y Neptuno) retienen sus atmósferas originales, pero los planetas terrestres (Mercurio, Venus, la Tierra y Marte) no lo hacen. Las actuales atmósferas de Venus, la Tierra y Marte (Mercurio carece de ella) se han ido constituyendo, sobre todo, a partir de escapes volcánicos de gases a lo largo de la historia del planeta. Las atmósferas originarias constaban, principalmente, de hidrógeno y helio, con menores cantidades de metano, amoniaco, nitrógeno y neon; las actuales son distintas. (Tenemos así que la terrestre se enriqueció en oxígeno, resultante del metabolismo de los organismos.)

¿Qué se hizo de las atmósferas originales de los planetas terrestres? Prevalece la hipótesis de que las atmósferas primitivas las barrió violentamente un huracán de materia expulsado por el Sol. La naturaleza del viento que barrió el sistema solar interno nunca se llegó a desentrañar del todo. Por los datos que disponemos, diríase que los planetas terrestres estuvieron sometidos a la misma clase de violenta efusión de gas que se ha observado en la nube molecular de Orión, AFGL 490 y L1551.

Podemos apoyarnos en las efusiones de energía para explicar la intrigante longevidad de las nubes moleculares gigantes. Cuando se observaron las primeras anchuras de líneas espectrales de 1 a 10 kilómetros por segundo en grandes regiones de nubes moleculares, se creyó que las velocidades, demasiado altas para responsabilizar a la agitación térmica, tendrían su razón de ser en el caso de que las nubes estuvieran colapsando bajo la influencia de la gravitación, lo que parecía plausible puesto que éstas son sede activa de formación de estrellas. La objeción que surge ante esa afirmación es que las nubes hubieran colapsado por entero en un millón



FLUJO DE MONOXIDO DE CARBONO de dos lobulos, observado en la vecindad de la fuente infrarroja AFGL 490, que está inmersa en una nube molecular gigante de unos 300 años-luz de diámetro en la constelación de la Jirafa. La parte central del gas que está corrida hacia el rojo, o alejándose (*curvas en color*), parece parcialmente descompensada con la parte central del gas que está corrido hacia el azul, o acercándose (*curvas en negro*). Esto sugiere que se está expulsando el gas en dos corrientes directamente opuestas, a partir de la vecindad de AFGL 490, indicada por la cruz en esta fotografía tomada con el telescopio Schmidt de 1,2 metros de Monte Palomar. La región de la emisión del monóxido de carbono con doble lóbulo tiene un diámetro aproximado de un año-luz. Los cuatro perfiles de la derecha recogen la emisión del monóxido de carbono en cuatro puntos del flujo molecular bipolar o cerca de él. Las máximas velocidades del gas se encuentran en la posición 3, junto a la posición de la propia AFGL 490. Las observaciones fueron realizadas por Paul M. Harvey, de la Universidad de Texas en Austin, y el autor con el radiotelescopio de 11 metros del Observatorio Nacional de Kitt Peak.

aproximado de años, siendo así que hay sólidos fundamentos para defender que las nubes tienen vidas que duran desde 10 hasta 100 millones de años. Además, si las nubes moleculares se encontraran todas en estado de colapso, la velocidad de formación de estrellas en la galaxia rebasaría, con mucho, la que de hecho se observa. Algo ha estado impidiendo que las nubes moleculares colapsen, y al hacerlo, ha inyectado suficiente energía mecánica en el gas para mantener anchuras de línea de varios kilómetros por segundo en toda la extensión de la nube.

Con el descubrimiento de los chorros moleculares de alta velocidad en torno a jóvenes estrellas de nubes oscuras, puede resolverse el misterio. Si cada estrella *B* da lugar, en el curso de la primitiva evolución, a la clase de chorros moleculares observados en la nube molecular de Orión de AFGL 490 la velocidad registrada de formación de estrellas *B* en los subgrupos de asociaciones *OB* puede ser adecuada para justificar la energía mecánica que mantiene 100.000 masas solares de gas molecular moviéndose a velocidades medias de dos kilómetros por segundo durante un prolongado lapso temporal, de hasta 10 millones de años.

Discos circumestelares

Aunque el descubrimiento de los chorros de alta velocidad en las nubes moleculares parece acabar con varios problemas que vienen de antaño, plantea otros nuevos e igualmente enigmáticos. La extraordinaria energía de los chorros y la alta frecuencia con que se dan no se reconcilian de buenas a primeras con los puntos de vista en vigor sobre la formación y primitiva evolución de las estrellas. Las nuevas pruebas de que hay procesos físicos de alta energía que acompañan a la creación de estrellas de mayor masa que el Sol (y tal vez, a la de estrellas de masa no superior) no pueden predecirse ni explicarse en el marco de las teorías existentes. ¿Qué mecanismo subyacería a las efusiones de alta velocidad que rodean a las estrellas jóvenes?

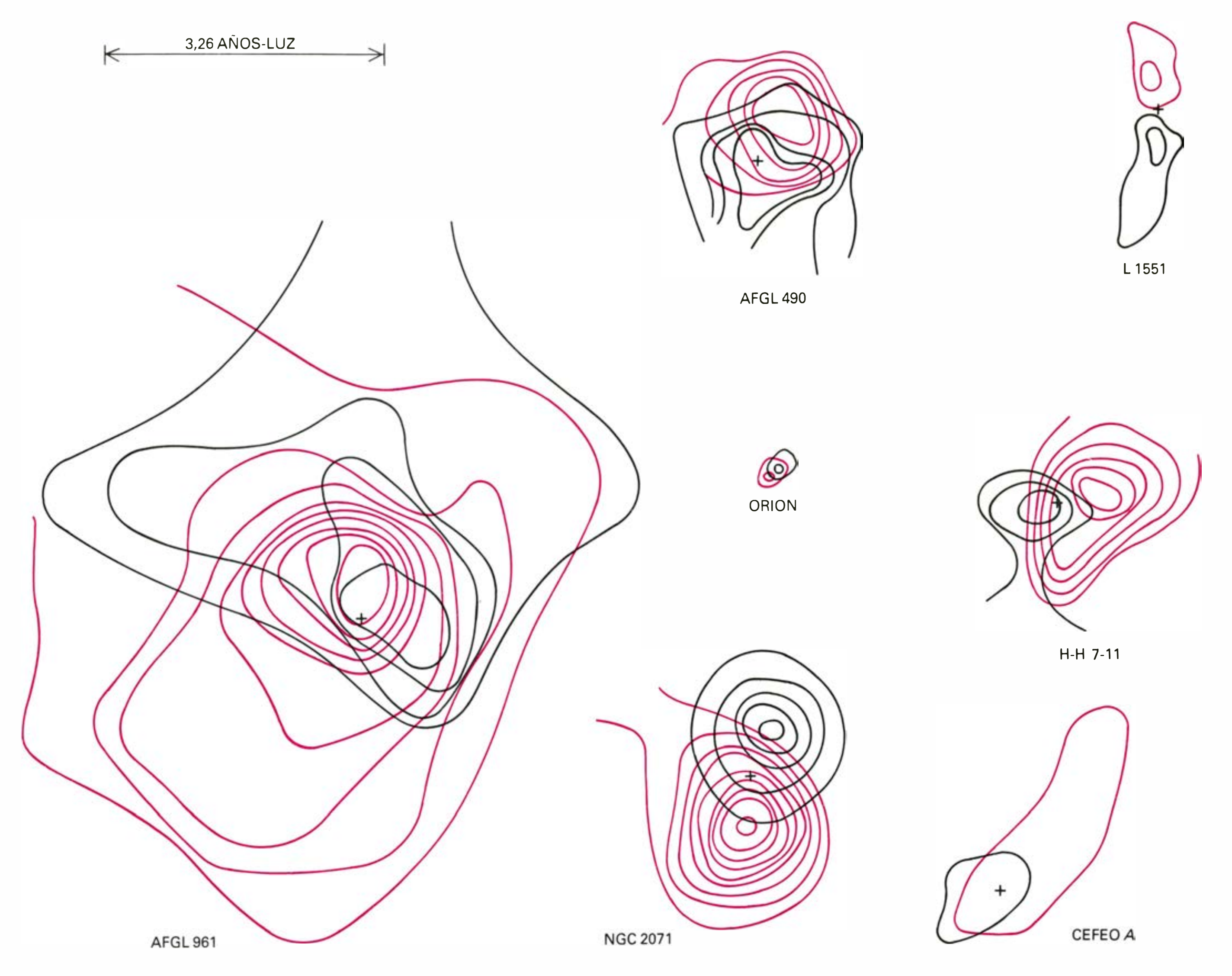
Cualquier explicación que se avance habrá de dar cuenta de dos propiedades importantes: la enorme energía involucrada en las efusiones y, en varios ejemplos fundamentales, la colimación, o focalización, de las corrientes para formar chorros bipolares. No es necesario suponer que el grueso de la materia que fluye haya sido expulsada de la fuente central. Con toda probabilidad, no ha sido expulsada. Más bien parece que la corriente esté formada por gases de la nube molecular arrastrados por un pequeño volumen de materia expulsado desde la estrella joven.

Quizá se explique mejor la distribución bipolar o en doble lóbulo, de la corriente que la fuente de energía. Si, por ejemplo, la corriente saliente fuera inicialmente esférica, podría canalizarse en una corriente bipolar en el caso de que el objeto central estuviera inmerso en un disco de gas. El gas saliente encontraría mucha menos resistencia al fluir a través de los polos del disco que al atravesar el ecuador; en virtud de lo cual, la corriente se canalizaría en dos sentidos opuestos.

¿Hay alguna prueba que indique que tales discos podrían formarse alrededor de estrellas jóvenes y protoestrellas? Hay numerosas aportaciones indirectas. Nuestro propio Sol constituye un ejemplo. De acuerdo con la opinión dominante, los planetas y el Sol se formaron al mismo tiempo a partir de la misma nube protoestelar en colapso. Evidentemente, la nube en colapso tomó la forma de un delgado disco, de suerte que, cuando los planetas acabaron por condensarse, terminaron todos ellos más o menos en el mismo plano, muy próximo al plano ecuatorial del Sol, donde permanecen hoy en órbita. En una escala inmensamente mayor, la nube en colapso que engendró la propia galaxia tomó la forma de un delgado disco (como lo revela la estrechez de la Vía Láctea si la miramos en una noche despejada y limpia). La verdad es que los cálculos teóricos muestran que, cuando una esfera giratoria de gas colapsa, origina inevitablemente un disco en rotación.

Todavía está fresca la noticia que da cuenta de la comprobación observacio-

DOS CHORROS DE GAS dirigidos en sentido opuesto, limpiamente separados en la vecindad de la fuente infrarroja oculta L1551 en la constelación Tauro, marcada por una cruz cerca del centro. La fotografía se sacó con el telescopio Schmidt de 1,2 metros de Monte Palomar. Como en AFGL 490, las corrientes proceden del objeto infrarrojo, del que se presume se trate de una estrella joven. La corriente corrida hacia el azul (*curvas negras*) parece incluir tres manchas nebulosas de emisión visible designadas como objetos Herbig-Haro: H-H 28, H-H 29 y H-H 102. Los dos primeros se mueven con una rapidez tal (150 kilómetros por segundo) que sus posiciones han cambiado en fotografías hechas con sólo unos años de intervalo. Las flechas indican la dirección de su movimiento. Extrapolando sus movimientos hacia atrás, se halla que los objetos fueron expulsados desde una región próxima a L1551 hace sólo 3000 años. El mapa de la emisión del monóxido de carbono fue construido por R. L. Snell, de Massachusetts, Robert B. Loren, de la Universidad de Texas en Austin, y Richard L. Plambeck, de la de California en Berkeley.

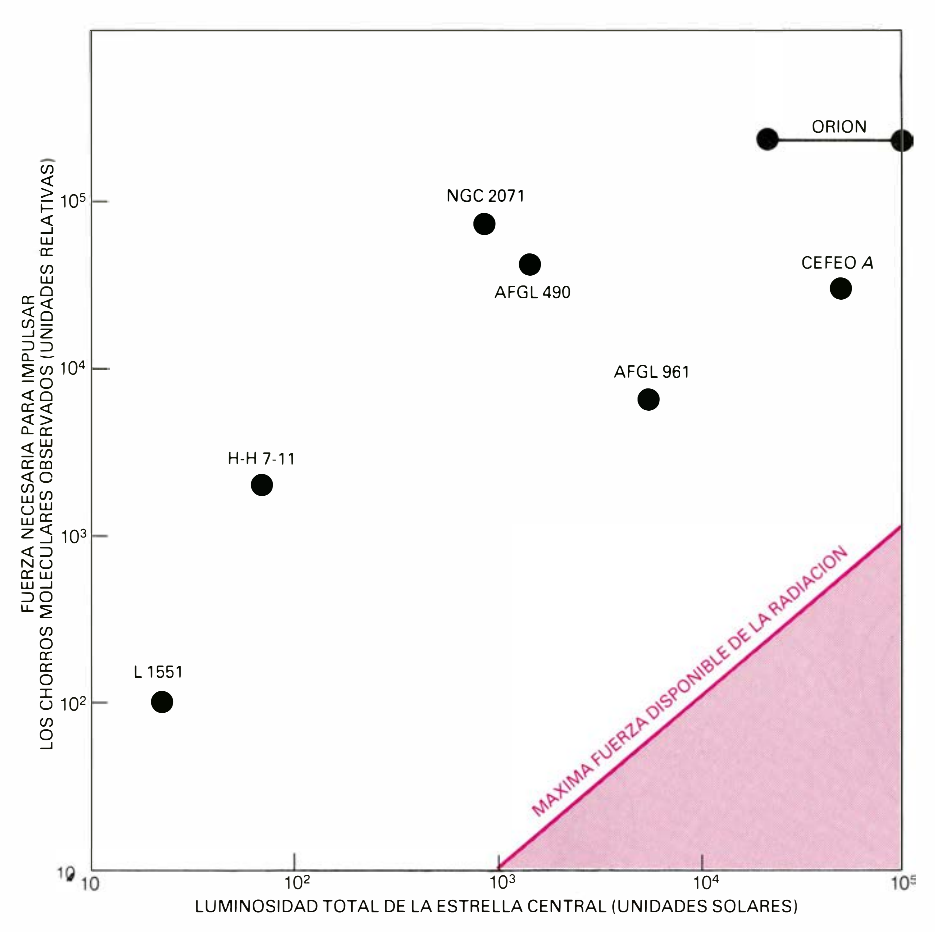


SIETE SISTEMAS DE CHORROS de alta velocidad alrededor de siete jóvenes objetos estelares, representados a la misma escala. Las líneas en color muestran emisiones de monóxido de carbono con corrimiento hacia el rojo; en negro presentan un corrimiento hacia el azul. La emisión de monóxido de carbono en cada fuente se origina en dos lóbulos centrados en un objeto infrarrojo (*cruz*). El mapa de Orión fue construido por Erickson, Goldsmith, Huguenin, Snell, Ulich y el autor con el Telescopio de Espejo Múltiple. Los de AFGL 490 y AFGL 961 fueron hechos por Harvey, T. N. Gautier, de la Universidad de Arizona, y el autor. El mapa de NGC 2071 lo realizó John Bally, de los Laboratorios Bell. Snell y Suzan Edwards, del Smith College, cartografiaron H-H 7-11. El mapa de Cefeo A se debe a Luis F. Rodríguez, del Instituto Mexicano de Astronomía, Paul T. P. Ho, de Berkeley, y James M. Moran, del Centro de Astronomía del Observatorio del Harvard College y del Smithsoniano. En el epígrafe precedente se citan los autores del mapa de L 1551.

nal de la existencia de un disco alrededor de una estrella con un peculiar espectro infrarrojo, la estrella MWC 349 en la constelación del Cisne. Para Rodger I. Thompson y Peter A. Strittmatter, de la Universidad de Arizona, y Edwin F. Erickson, Fred C. Witteborn y Donald W. Strecker, del Centro de Investigación Ames de la Administración Nacional de Aeronáutica y del Espacio, el espectro de la estrella está producido, en forma muy plausible, por un disco circumestelar. Todavía están por realizar estudios exhaustivos de las líneas espectrales en el infrarrojo causadas por las estrellas jóvenes ocultas asociadas con cualquiera de los chorros moleculares; pero los datos disponibles para dos fuentes infrarrojas, AFGL 490 y AFGL 961, muy afines, sugieren la presencia de un disco alrededor del objeto central. Los resultados obtenidos hasta la fecha, aunque provisionales, apuntalan la creencia de que las estrellas más recientes están rodeadas de un disco de gas y polvo.

Naturaleza de la fuente de energía

Pero el tema se enturbia cuando se trata de la fuente de energía suficiente para mantener los dos chorros de alta velocidad. En el caso de AFGL 490, un objeto central de no más de 15 masas solares impulsa 30 masas solares de gas molecular que fluye hacia fuera. En otras palabras, cualquiera que sea el proceso productor de energía, acumula la necesaria para impulsar y acelerar hasta grandes velocidades una masa de materia que viene a duplicar la masa central. Basándose en las observaciones en el infrarrojo lejano realizadas desde una altura de 12.000 metros por el Observatorio Volante Kuiper, de la NASA, Harvey, Murray F. Campbell y William F. Hoffmann, de la Universidad de Arizona, han calculado que AFGL 490 tiene una luminosidad total de 1400 veces de solar. Por consiguiente, AFGL 490 tardaría 1000 años en emitir suficiente energía luminosa para igualar la cantidad de energía mecánica transportada por el gas molecular saliente. Si la corriente gaseosa ha estado fluyendo 10.000 años, como se ha estimado, se deduce que el 10 por ciento de la energía luminosa emitida por el objeto central se ha convertido en energía mecánica. O bien es así, o bien la



INADECUACION DE LA PRESION DE RADIACION para acelerar los siete sistemas de chorros moleculares cartografiados en la ilustración de la página anterior. La masa de gas molecular de una corriente bipolar puede frecuentemente exceder la de 10 soles. Para acelerar tamaña masa hasta las altas velocidades observadas en los chorros moleculares se necesita una fuerza enorme. La línea diagonal indica la máxima fuerza que puede generarse por la presión de radiación disponible en los jóvenes objetos estelares del centro de los chorros. Incluso para estrellas que emitan 10.000 veces la radiación solar, o 4 × 10^{37} erg por segundo, la presión de radiación resulta insuficiente, en muchos órdenes de magnitud, para proporcionar la fuerza requerida. El diagrama está basado en datos compilados por John Bally y el autor.

energía mecánica la está suministrando otra fuente que actúa en paralelo con la fuente radiante.

Se agrupan en dos categorías las demás fuentes de energía posibles: un proceso que lleva consigo una o más expulsiones explosivas de gas desde el objeto central y un proceso que da por resultado una salida continua de materia gaseosa desde éste. Si los chorros se originaron en un único suceso explosivo, la energía original de la explosión tendría que haber sido mucho mayor que la energía cinética observada hoy en la corriente molecular de alta velocidad. La razón para ello es que la mayor parte de la energía inicial de la explosión habría sido radiada hacia fuera en ondas de choque generadas por la materia eyectada. Sólo una fracción de la energía podría haberse depositado directamente en forma de energía mecánica. Pese a ello, una explosión de supernova oculta en una nube molecular podría fácilmente liberar energía suficiente para explicar la corriente hacia fuera. El punto débil de esta hipótesis estriba en que la tasa de nacimiento de supernovas en la vecindad del sistema solar es unas 20 veces menor que la tasa de nacimiento de chorros moleculares bipolares.

Las pruebas de observación proporcionan un argumento más directo contra la hipótesis de una explosión única. Con el radiotelescopio de 11 metros de Kitt Peak, el autor ha realizado observaciones de alta sensibilidad de los chorros salientes en AFGL 490 y ha hallado que el gas de velocidad mayor cubre solo una fracción pequeña del área total del gas en movimiento. Significa ello que la corriente de alta velocidad está decelerándose fuertemente a medida que se aleja de la fuente. Sería de esperar que también el material expulsado por una explosión única se decelerase al chocar con la nube circundante; en tal caso, no se percibiría variación de velocidad a través del chorro, porque toda la materia expulsada en una explosión única se deceleraría al mismo tiempo y tendría hoy la misma velocidad. Otra objeción a la hipótesis de explosión única: un gran brote de energía solitario podría destrozar el disco circumestelar necesario para mantener el chorro en forma bipolar.

Fuentes continuas de energía

Con semejantes razones, habremos de volver la vista hacia una corriente continua o una serie de brotes explosivos por si explican mejor las observaciones. ¿Qué ocurre? Aunque no se puede descartar una sucesión de pequeñas explosiones, la expulsión continua parece más probable. Hay precedentes de este fenómeno en astrofísica. Se sabe que la mayoría de las estrellas, incluido el Sol, están perdiendo masa en forma de viento estelar. Para el caso del Sol, su velocidad actual de pérdida es muy pequeña: en torno a un 10^{-14} de su masa total por año. La cantidad de movimiento y la energía en un viento tan débil son, con mucho, demasiado pequeñas para impulsar las corrientes de salida observadas en los chorros moleculares. Parece probable, sin embargo, que cuando el Sol era joven, su velocidad de pérdida de masa fuese mucho mayor, quizás hasta de 10^{-7} de su masa por año. Las estrellas T de Tauro, objetos que se cree son los progenitores de estrellas del tipo solar, están, según las observaciones, perdiendo masa a esa velocidad. Aun así, la cantidad de movimiento que tal viento imparte al gas molecular circundante sería todavía una décima parte de la cantidad de movimiento necesaria para explicar la corriente de salida alrededor de L1551 y una diezmilésima parte de la necesaria para explicar el flujo alrededor de AFGL 490 y la nube molecular de Orión.

Queda por considerar una fuente final de energía: la presión hacia fuera que ejerce la propia radiación estelar. Ya hemos visto que la energía cinética total transportada por los chorros moleculares es igual a no más del 10 por ciento de la energía luminosa emitida por las estrellas centrales a lo largo de la vida dinámica de los chorros. ¿Podría entonces la radiación ser la fuerza que alimenta los chorros?

Para evaluar la eficacia de la radiación como fuerza impulsora debemos calcular la presión en un campo de radiación: la cantidad de movimiento real transportada por los fotones. En un choque entre un fotón y una partícula de gas o polvo, el fotón proporciona cantidad de movimiento a la partícula. De acuerdo con la teoría especial de la relatividad, la cantidad de movimiento p transportada por un fotón está rela-

cionada con la energía E del fotón mediante la ecuación $p = E/c$, donde c es la velocidad de la luz. La luminosidad total de una estrella, L, es igual a la cantidad de energía emitida en fotones por segundo. Por tanto, si cada fotón es absorbido al menos una vez, la cantidad total de movimiento que la estrella puede impartir al gas y polvo circundantes por segundo es precisamente L/c. Por ejemplo, en el caso de las estrellas gigantes O, muy calientes, la cantidad de movimiento que pueden impartir los fotones excede a la cantidad de movimiento transportada por las partículas en los vientos estelares. ¿Bastaría una presión de radiación tan intensa para explicar los chorros moleculares?

John Bally, de los Laboratorios Bell, y el autor investigaron recientemente esta posibilidad. Compararon la fuerza debida a la presión de radiación con la fuerza necesaria para impulsar las corrientes de salida en el caso de los chorros moleculares mejor estudiados. En todos los casos, encontramos que la cantidad de movimiento del gas saliente era de 100 a 1000 veces mayor que la cantidad de movimiento disponible de la presión de radiación en la hipótesis de que cada fotón se absorbiese en sólo una ocasión. No parece, pues, que la presión de radiación impulse las corrientes moleculares observadas.

Si ni la liberación explosiva de energía ni la presión de radiación sirven para explicar los chorros moleculares que rodean las estrellas jóvenes, ¿cuál será el mecanismo impulsor? Esta pregunta fundamental va a mantener ocupados a los astrofísicos teóricos durante algún tiempo, sin duda. Cualquiera que sea la respuesta, no sólo nos desvelará los misterios persistentes entorno a la formación y primitiva evolución de las estrellas, sino que nos llevará también a la comprensión de los procesos que crean esos chorros enigmáticos hallados cerca de los quasars y radiogalaxias [véase "Chorros cósmicos", por Roger D. Blandford, Mitchell C. Begelman y Martin J. Rees; INVESTIGACIWN Y CIENCIA, julio). Las observaciones en el infrarrojo y en onda milimétrica han colocado en su sitio otra pequeña pieza del rompecabezas de la formación y evolución de las estrellas. ¿Qué otras piezas quedan por encontrar? ¿Cuántas más se necesitarán para resolver todo el rompecabezas? He ahí los emocionantes problemas del presente y el futuro, problemas que se abordarán y aclararán merced al progreso continuado en el frente observacional y en el frente teórico de la astrofísica.

Topoisomerasas de ADN

Por tales se entiende los enzimas que modifican la topología de los anillos de ADN, ensartando y anudando varios de ellos. Cabe presumir que intervengan, asismismo, en los procesos más fundamentales de la transmisión hereditaria

James C. Wang

La información que encierra una molécula de ADN está inscrita en su secuencia de pares de bases nucleotídicas complementarias. No depende de cuán retorcida, enredada o anudada se halle la molécula. Lo que no empece que a lo largo de los últimos diez años se haya descubierto que la forma topológica de una molécula de ADN ejerce una importante influencia sobre la actividad de ésta en la célula. Diríase que los enzimas topoisomerasas, que transforman el ADN de una configuración topológica en otra, desempeñan un papel fundamental en los episodios genéticos centrales de replicación, transcripción y recombinación del ADN.

La imagen asociada a la de una molécula de ADN es la de dos cadenas lineales retorcidas formando una hélice doble con un eje recto. Pero en muchos casos, se presenta como un anillo cerrado formado por una sola cadena o por dos cadenas retorcidas en doble hélice. Además, ambos tipos anulares pueden adoptar otras configuraciones topológicas. El de cadena simple puede anudarse. Dos anillos de cadena simple pueden unirse y retorcerse para forjar un anillo de doble cadena. Si se corta una de las cadenas de un anillo de dos, la doble hélice puede desenrollarse unas cuantas vueltas. Al desenrollarse, la molécula sufre una tensión, que provoca que el propio eje de la doble hélice se retuerza y forme a su vez una hélice. Cuando adopta esta estructura, se dice que el ADN está superenrollado.

Podría dar la impresión de que, para lograr tales conversiones topológicas, serían precisas muy distintas manipulaciones de las cadenas. La verdad es que todas esas transformaciones se basan en un mismo mecanismo general: el anillo debe romperse transitoriamente y un segmento del ADN atravesar la brecha antes de que los dos cabos se vuelvan a juntar. En 1971 se descubrió que bastaba una sola molécula de topoisomerasa para llevar a cabo la operación en todas sus partes de rotura, paso y cierre de la brecha. Pero hoy sabemos que hay dos clases de topoisomerasas. Las de una clase cortan una sola cadena de ADN. Las de la otra cortan las dos cadenas a la vez.

Desde 1971 se han venido descubriendo muchas topoisomerasas, de ambas clases, en organismos tan dispares como las bacterias y el hombre. A pesar de su ubicuidad, las topoisomerasas habrían permanecido como mera curiosidad (unas moléculas biológicas que realizan operaciones matemáticas) de no haberse observado que su manipulación del ADN coincidía con los momentos más importantes del ciclo genético de la célula. Para que la información genética de la molécula de ADN pueda replicarse, transcribirse en ARN o recombinarse con la de otros genes, deben acceder a las bases nucleotídicas gran variedad de enzimas. En algunas configuraciones topológicas, los pares de bases resultan accesibles a tales enzimas, pero en otras configuraciones no ocurre así. Por ejemplo, el superenrollamiento hace más fácil que se desenrolle la doble hélice de un anillo de doble cadena. Cuando se desenrollan las bases de la médula de la hélice, quedan expuestas a la acción de los enzimas. Parece que para algunos organismos el desenrollamiento es condición previa para la replicación del ADN y para la transcripción del ADN en ARN. Una topoisomerasa puede obligar a que el anillo de ADN adquiera la forma superenrollada; y otra puede más tarde liberarlo de semejante configuración. Entra, pues, dentro de lo verosímil que las topoisomerasas ayuden a controlar la tasa de replicación y transcripción a través de su regulación del grado de superenrollamiento del ADN.

Tras el descubrimiento de las moléculas circulares de ADN se comenzó a sospechar que la forma topológica del ADN influía sobre su estructura y función. A comienzos de los años 60, Robert L. Sinsheimer y Walter Fiers, del Instituto de Tecnología de California, encontraron que el material genético del virus ϕX174, que infecta bacterias, se presentaba en forma anular y con una sola cadena de ADN. El desarrollo de métodos eficaces para la determinación de la secuencia de bases de una molécula de ADN permitió examinar el anillo minuciosamente. Se pudo comprobar que el anillo carecía de interrupciones.

Se le reconoce a John Cairns, de la Universidad Nacional de Australia, el descubrimiento, en 1963, de anillos de ADN de doble cadena. Sabemos ya que el ADN de muchos organismos se presenta en forma de anillo de doble cadena. Tales anillos son normales en bacterias y virus. También se encuentran en la mitocondria, el orgánulo productor de energía de las células de los organismos superiores, y en el cloroplasto, donde tiene lugar la fotosíntesis de las células vegetales.

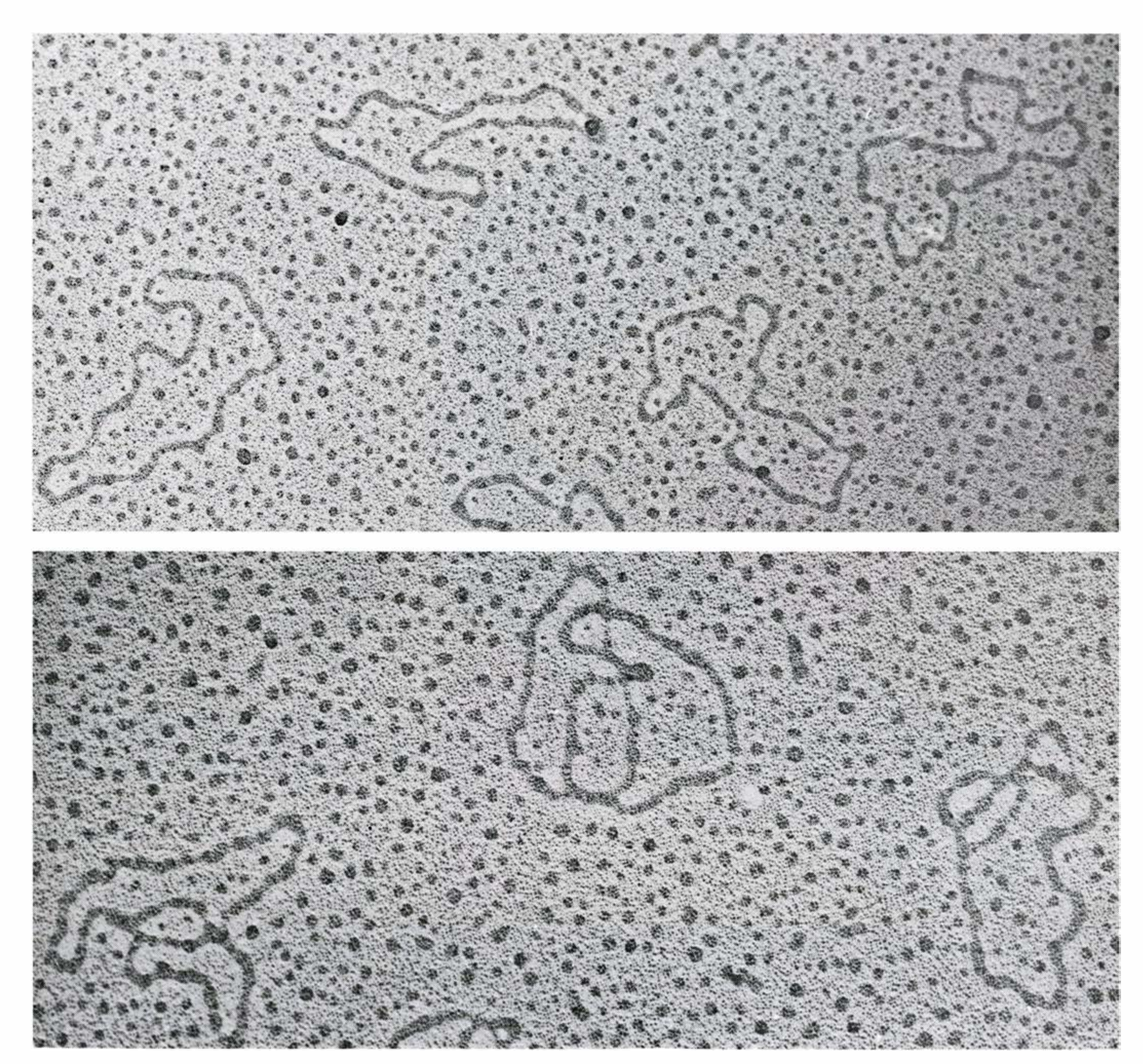
La configuración topológica define un nivel adicional de estructura en los anillos de cadena simple o doble. Son especialmente curiosas las configuraciones que adopta el anillo de doble cadena mientras mantiene intactas ambas cadenas. Una magnitud que influye de manera importante en la configuración del anillo de doble cadena es el número de veces que una cadena rodea completamente la otra. Esta magnitud se conoce con el nombre de índice de enlace. [Para una definición matemática más precisa del índice de enlace véase "ADN superenrollado" por William R. Bauer, F. H. C. Crick y James H. White; INVESTIGACIÓN Y CIENCIA, septiembre de 1980.]

El índice de enlace es una propiedad topológica del anillo de doble cadena. Las propiedades topológicas dependen de la disposición tridimensional de una estructura. Son, sin embargo, completamente diferentes de las propiedades geométricas. El principal carácter de

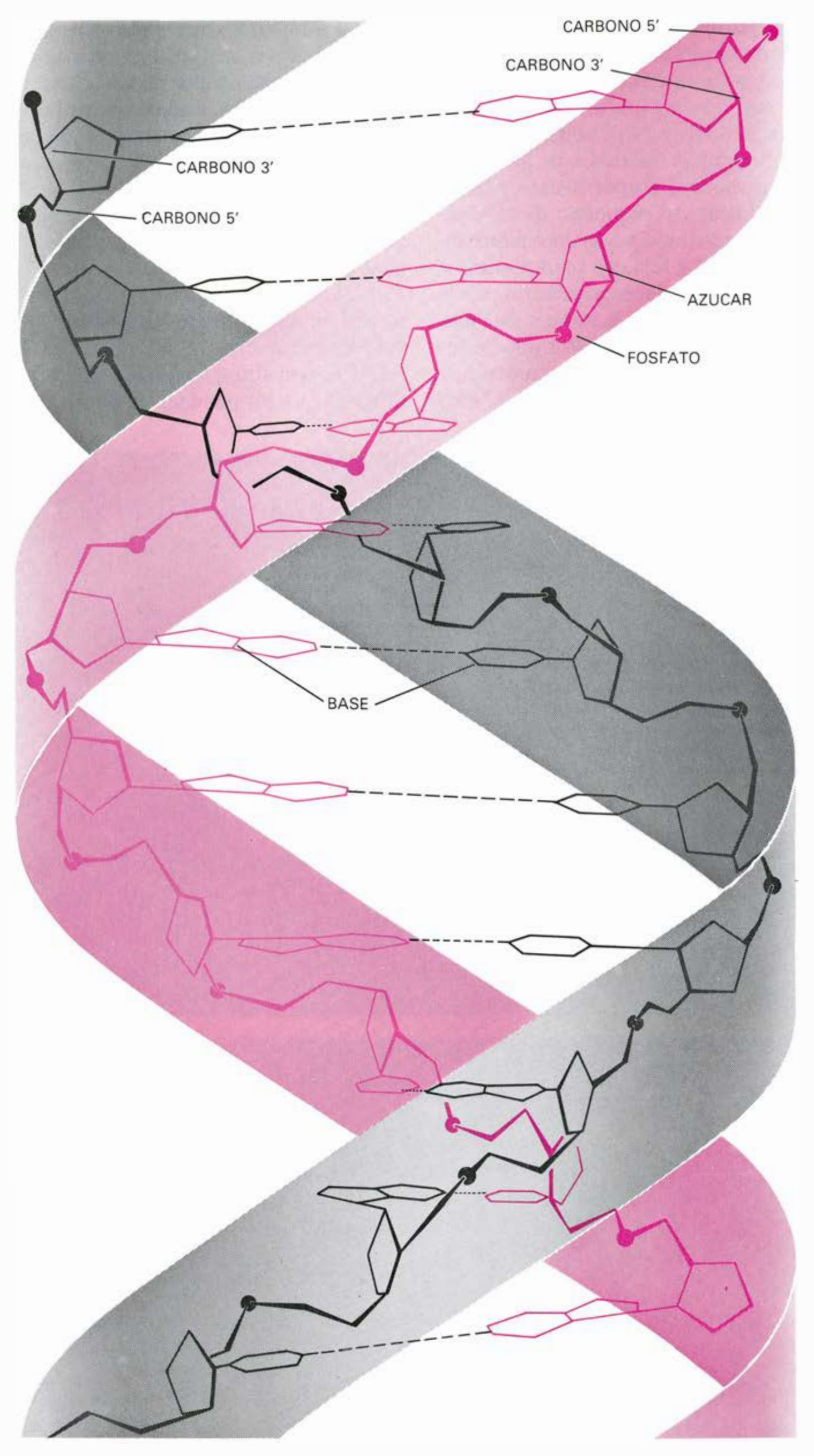
una propiedad topológica es que permanece inalterada cuando cambia la forma de la estructura. No importa cuánto se estire o retuerza el anillo de doble cadena, su índice de enlace no cambiará mientras persistan enteras ambas cadenas.

Para entender cómo influye el índice de enlace en la configuración y propiedades físicas de un anillo conviene considerar la estructura de la molécula de ADN. Las moléculas biológicas poseen varios niveles estructurales, comenzando por el nivel de los átomos y los enlaces químicos covalentes que los unen. (Se forma un enlace covalente cuando dos átomos adyacentes comparten electrones; se trata del tipo de enlace químico más fuerte.) Los átomos de la molécula de ADN están dispuestos en tres tipos de grupos: unidades de azúcar de cinco carbonos, grupos fosfato y bases nucleotídicas. El esqueleto de la cadena está constituido por grupos alternantes de azúcar y fosfato, unidos entre sí por medio de enlaces covalentes. Cada unidad de azúcar de la cadena está unida al grupo fosfato situado a uno de sus lados por el átomo de carbono designado 5' y, al grupo fosfato situado al otro lado, por el átomo de carbono designado 3'. Por tanto, el esqueleto de azúcar-fosfato está polarizado: puede recorrerse en dos sentidos opuestos, según se pase del carbono 5' al carbono 3' de cada molécula de azúcar o se recorra el camino inverso.

A cada molécula de azúcar se engarza una base: adenina, timina, citosina o guanina, que se simbolizan por *A*, *T*, *C*, *G*. Las bases sobresalen del esqueleto casi en ángulo recto. El segundo nivel de estructura en la molécula de ADN lo constituyen los enlaces de hidrógeno que forman entre sí las parejas



SE CUENTA LA FORMACION DE NUDOS en anillos de ADN entre las virtualidades de los enzimas denominados topoisomerasas. El material genético de muchos organismos tiene la forma de un anillo, constituido por una cadena de ADN, o dos cadenas entrelazadas en una doble hélice. El anillo puede adoptar múltiples configuraciones topológicas. El paso del anillo de una configuración a otra está catalizado por topoisomerasas. La micrografía electrónica superior muestra anillos de ADN de cadena simple de un virus, el bacteriófago *fd*, que infecta bacterias. La micrografía electrónica inferior muestra los anillos después de exponerlos a la topoisomerasa de la bacteria *Escherichia coli*. Cortando las cadenas, pasando un segmento del anillo por la brecha y volviendo a unir los cabos el enzima ha hecho un nudo en cada anillo [*véase las ilustraciones de las páginas 66 y 67*]. El anudamiento sólo se ha observado hasta ahora en el laboratorio. El proceso de corte, paso y reunión es esencial para la acción de todas las topoisomerasas. Algunos de los enzimas, designados como del tipo I, cortan una cadena simple de ADN. Otros, designados del tipo II, pueden cortar ambas cadenas de la doble hélice.



DOBLE HELICE DE ADN, que consta de dos cadenas entrelazadas en sentido dextrógiro. Los esqueletos de ambas cadenas se sitúan en la parte exterior de la hélice; cada uno de ellos es una cadena en que alternan unidades de azúcar y de fosfato unidas por enlaces químicos fuertes, del tipo denominado enlace covalente. Cada molécula de azúcar de cinco carbonos está unida al grupo fosfato de un lado a través del átomo de carbono designado 5' y, al grupo fosfato del otro, por el átomo de carbono designado como 3'. Por tanto, el esqueleto tiene polaridad. Se puede recorrer en toda su extensión en el sentido que va del carbono 5' al carbono 3' de cada azúcar, o en sentido puesto. Las dos cadenas avanzan en sentidos contrarios. Unida a cada azúcar y proyectándose hacia el interior de la hélice hay una de las cuatro bases adenina (*A*), timina (*T*), citosina (*C*) o guanina (*G*). Las bases de las cadenas opuestas se mantienen unidas a través de enlaces de hidrógeno, que son más débiles que los enlaces covalentes. *A* sólo puede formar enlaces con *T* y *C* sólo puede formarlos con *G*. La secuencia de pares de bases constituye la información genética que lleva la molécula. La forma y composición de los residuos del ADN determinan que en la célula viva la hélice sea más estable cuando las dos cadenas giran una alrededor de la otra dando una vuelta cada 10,5 pares de bases. La estructura que se muestra es la que se ha deducido de observaciones con fibras cristalinas de ADN, más estable cuando las cadenas dan una vuelta cada 10 pares de bases.

de bases, una en cada cadena de la doble hélice. (El enlace de hidrógeno se forma por atracción electrostática entre un átomo de hidrógeno, enlazado a un átomo de oxígeno o de nitrógeno, y otro átomo de oxígeno o nitrógeno.)

Si se juntan los esqueletos de azúcar-fosfato de dos cadenas con sus bases enlazadas, las bases de una cadena forman puentes de hidrógeno con las bases de la otra cadena. Dada la forma y composición química de las bases, *A* se enlaza sólo con *T* y *C* sólo puede hacerlo con *G*. De ahí que siempre que aparezca *A* en la primera cadena deba aparecer *T* en la segunda, y viceversa. Las mismas relaciones valen para *C* y *G*. Se dice que son complementarias las cadenas cuyas bases nucleotídicas se emparejan así. En una molécula de doble cadena, las cadenas complementarias progresan en sentidos opuestos: una de 5' a 3' y, la otra, de 3' a 5'.

La geometría y la composición química de los residuos del ADN determinan que, cuando se cierren los enlaces de hidrógeno entre las bases, las cadenas adopten la forma de doble hélice, con el esqueleto de azúcar-fosfato en la parte externa y las parejas de bases en la parte interna. La hélice es dextrógira, es decir, se enrosca en el mismo sentido que lo hace un tornillo ordinario. La estructura de doble hélice se caracterizó por primera vez en fibras cristalinas, donde se demostró que cada cadena daba una vuelta completa alrededor de la cadena complementaria cada 10 pares de bases. Es la estructura que se denomina hélice *B* de Watson y Crick. Resultados obtenidos recientemente por mi grupo, en la Universidad de Harvard, y por el de Aaron Klug, en el British Research Council Laboratory of Molecular Biology, de Cambridge, prueban que las cadenas de la doble hélice de la molécula de ADN en disolución dan una vuelta completa cada 10,5 pares de bases.

Como señalé antes, el valor del índice de enlace de un anillo de doble cadena es igual al número de veces que una cadena gira alrededor de la otra. Si una molécula lineal de ADN en que cada cadena gira una vez cada 10,5 pares de bases se cierra para formar un anillo, el valor de su índice de enlace será igual al número de pares de bases dividido por 10,5. Hasta el virus más pequeño tiene ADN con miles de pares de bases, de modo que el valor del índice de enlace para un anillo de doble cadena en tales organismos variará entre centenares y centenares de miles de unidades. Entre las operaciones topológicas que realizan las topoisomerasas

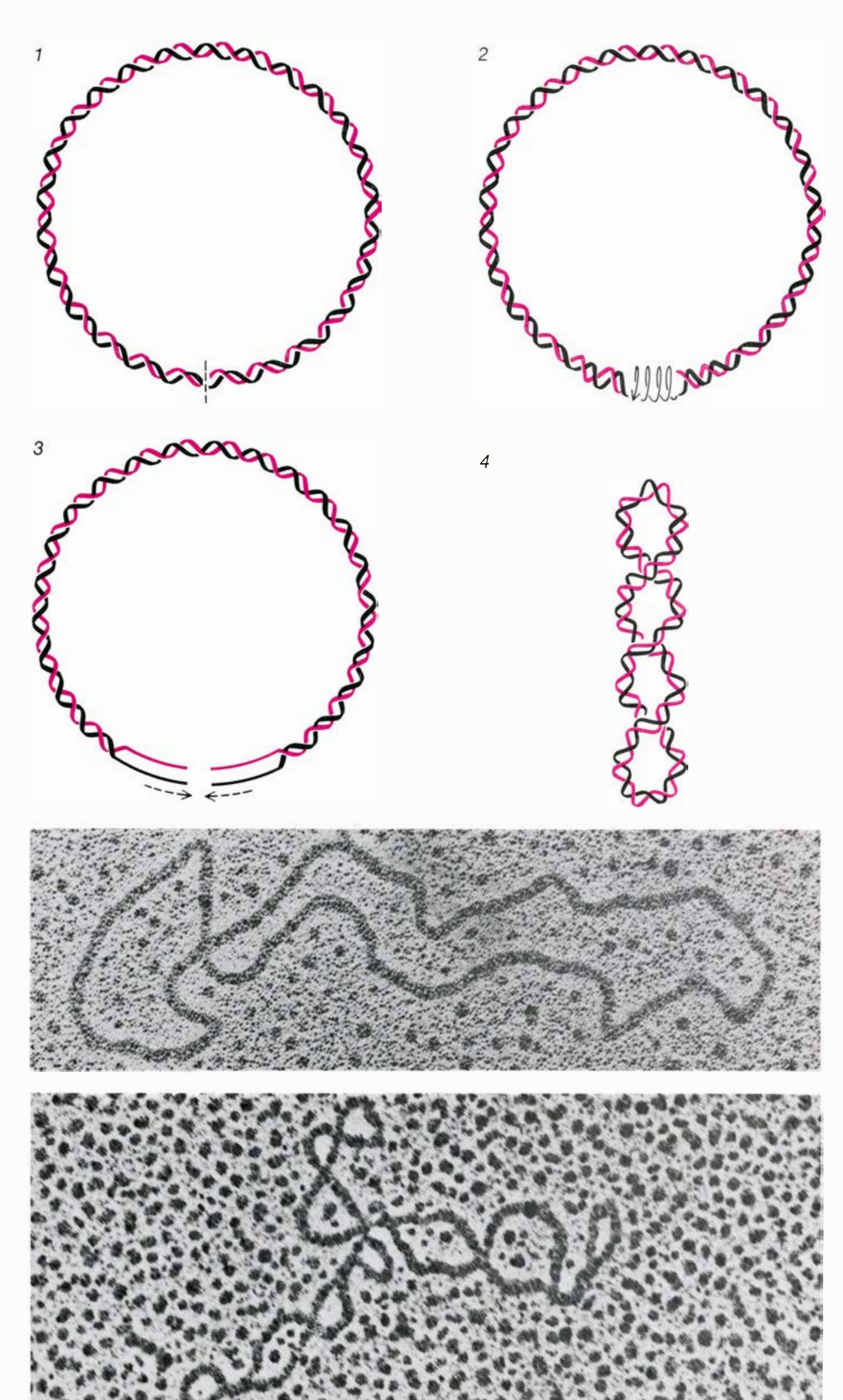
se cuenta, como es sabido, la alteración del valor del índice de enlace de un anillo de doble cadena.

Se consideran isómeros dos moléculas que tienen la misma composición química pero difieren en la estructura. Por tanto, dos anillos de doble cadena de ADN con la misma secuencia de pares de bases pero con valores diferentes del índice de enlace, serán isómeros. Al divergir en una magnitud topológica, se llaman topoisómeros. De ahí que el enzima que altera la forma topológica se llame topoisomerasa.

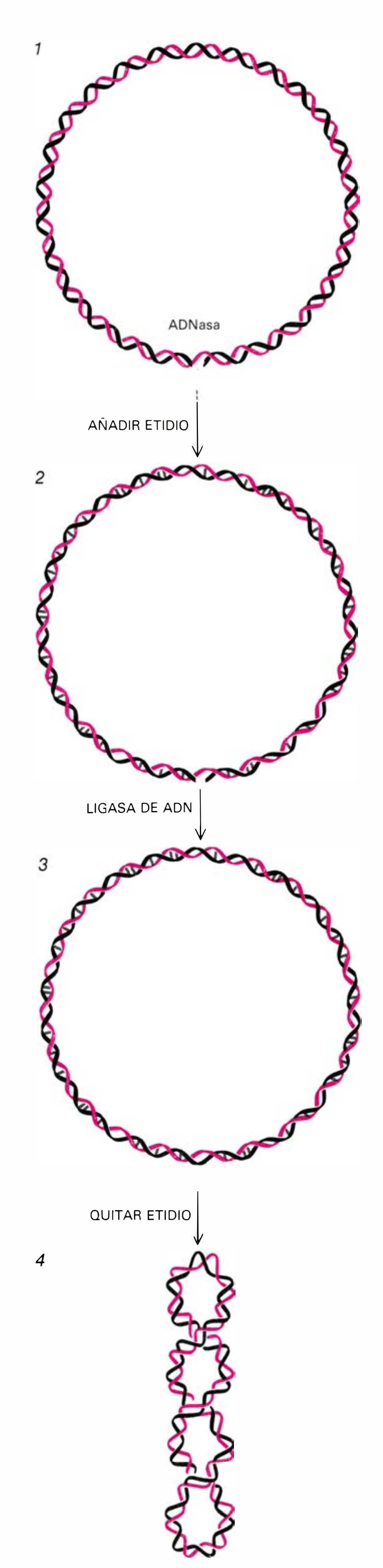
Cuando se reduce el valor del índice de enlace de un anillo de doble cadena, se induce la torsión y el enrollamiento de la doble hélice de ADN. En este estado se dice que el anillo está superenrollado. Jerome Vinograd, del Instituto de Tecnología de California, fue el primero en detectar anillos superenrollados. Vinograd y sus colegas observaron que el ADN del virus de polioma, que infecta animales, poseía propiedades físicas diferentes de las que cabía esperar atendiendo sólo a su tamaño. El ADN de ese virus era un anillo de doble cadena que constaba de unos 5000 pares de bases. Vinograd probó que las propiedades del anillo se debían al superenrollamiento provocado por una reducción en el valor del índice de enlace.

Puede demostrarse preparando una serie de anillos que difieran sólo en el valor del índice de enlace que el valor del índice de enlace de un anillo de ADN, y por tanto su grado de superenrollamiento, ejerce una profunda influencia sobre las propiedades físicas del anillo. Antes del descubrimiento de las topoisomerasas se conocía un método para preparar ese conjunto de moléculas. Se parte de una disolución de anillos de doble cadena de longitud conocida y cualquier valor del índice de enlace. Se emplea el enzima pancreático conocido con el nombre de ADNasa I para cortar una cadena en cada anillo. La ADNasa rompe el enlace covalente situado entre un grupo de azúcar y un grupo de fosfato. (Al mismo tiempo se escinde una molécula de agua, una de cuyas partes resultantes se une al grupo azúcar, mientras la otra lo hace al grupo fosfato.)

Con una de las cadenas rota, no existe restricción alguna sobre el valor del índice de enlace del anillo. La cadena cortada puede girar libremente alrededor de la cadena intacta y subir o bajar el valor del índice de enlace. En la molécula lineal, una cadena gira alrededor de la otra una vez cada 10,5 pares de bases, porque esa configuración es la que produce la mínima tensión sobre la doble hélice: se trata de la configura-



SUPERENROLLAMIENTO de un anillo de ADN de doble cadena, que modifica el anillo para conferirle un aspecto más retorcido y compacto. La forma del anillo de ADN está fuertemente afectada por el número de veces que una cadena gira alrededor de otra; al valor de ese número se le denomina índice de enlace. Es una magnitud topológica: no puede alterarse mientras las cadenas permanezcan intactas, sin que importe cómo se estire o retuerza el anillo. Sin embargo, si las cadenas se cortan (*1*) y luego se hace girar una alrededor de la otra en sentido opuesto al del enrollamiento de la hélice (*2*), la hélice se abre. Cuando los extremos cortados se vuelven a unir (*3*), el valor del índice de enlace ha disminuido en el número de vueltas que se han dado. (La ilustración muestra el fundamento de la reducción del valor del índice de enlace; las topoisomerasas no tienen que operar necesariamente por el mismo mecanismo.) Las cadenas del ADN en una molécula lineal giran una vez cada 10,5 pares de bases, porque esa configuración es la que menos tensión impone a la doble hélice. De un anillo con una relación de 10,5 se dice que está relajado. Al aumentar o disminuir el valor de la relación se introduce tensión en la doble hélice, a la que responde con el superenrollamiento (*4*). La disminución en el valor del índice de enlace produce superenrollamiento negativo; el aumento en el valor del índice de enlace conduce al superenrollamiento positivo. La micrografía superior muestra el anillo de ADN relajado de un virus bacteriano denominado *PM2*. En la micrografía inferior se presenta un anillo con superenrollamiento negativo de un mismo virus.



EL ETIDIO RELAJA LA TORSION de la hélice. Se trata de una molécula que se desliza entre pares de bases adyacentes cuando se une al ADN. Se puede utilizar para preparar una familia de anillos de ADN de doble cadena que difieren en el valor del índice de enlace y, por consiguiente, tienen distintos grados de superenrollamiento. El enzima denominado ADNasa corta una cadena del anillo (*1*), lo que permite a ésta girar alrededor de la cadena intacta y llevar al anillo a su estado de mínima tensión: el estado relajado. Se insertan entonces las moléculas de etidio en la hélice relajada (*2*). Cada molécula de etidio abre la hélice 26 grados, de modo que 14 moléculas reducen en una unidad el valor del índice de enlace. Después de la unión del etidio, se reúnen los cabos por medio de ligasa de ADN, con lo que el nuevo valor del índice de enlace queda sellado en el anillo (*3*). Cuando se elimina el etidio, el anillo se superenrolla negativamente, debido a la disminución en el valor de su índice de enlace (*4*). Se puede, por tanto, controlar el grado de superenrollamiento modificando la cantidad de etidio presente en la disolución. Los anillos de ADN de doble cadena que tienen la misma composición química, pero difieren en el valor del índice de enlace, presentan distintas formas y propiedades físicas; se les denomina por ello topoisómeros. De ahí el nombre de topoisomerasa que se aplica al enzima que cambia el valor del índice de enlace y altera, con ello, la configuración de los anillos.

ción más estable de la molécula. Cuando se corta una cadena, la doble hélice vuelve gradualmente a esta forma. Cuando el índice de enlace alcanza un valor igual al número de pares de bases dividido por 10,5, se dice que el anillo se ha relajado.

Podemos reducir el valor del índice de enlace por debajo del que corresponde al estado relajado empleando etidio, una molécula plana que se desliza entre dos pares de bases cuando se une a la doble hélice. Cada molécula de etidio enlazada al ADN desenrolla la hélice unos 26 grados. Por tanto, 14 moléculas de etidio desenrollarán la hélice una revolución completa, y rebajarán en una unidad el valor del índice de enlace. Del número de moléculas de etidio unidas depende la proporción en que se rebaje el valor del índice de enlace y, a su vez, el número de moléculas unidas dependerá de la concentración de etidio en la disolución.

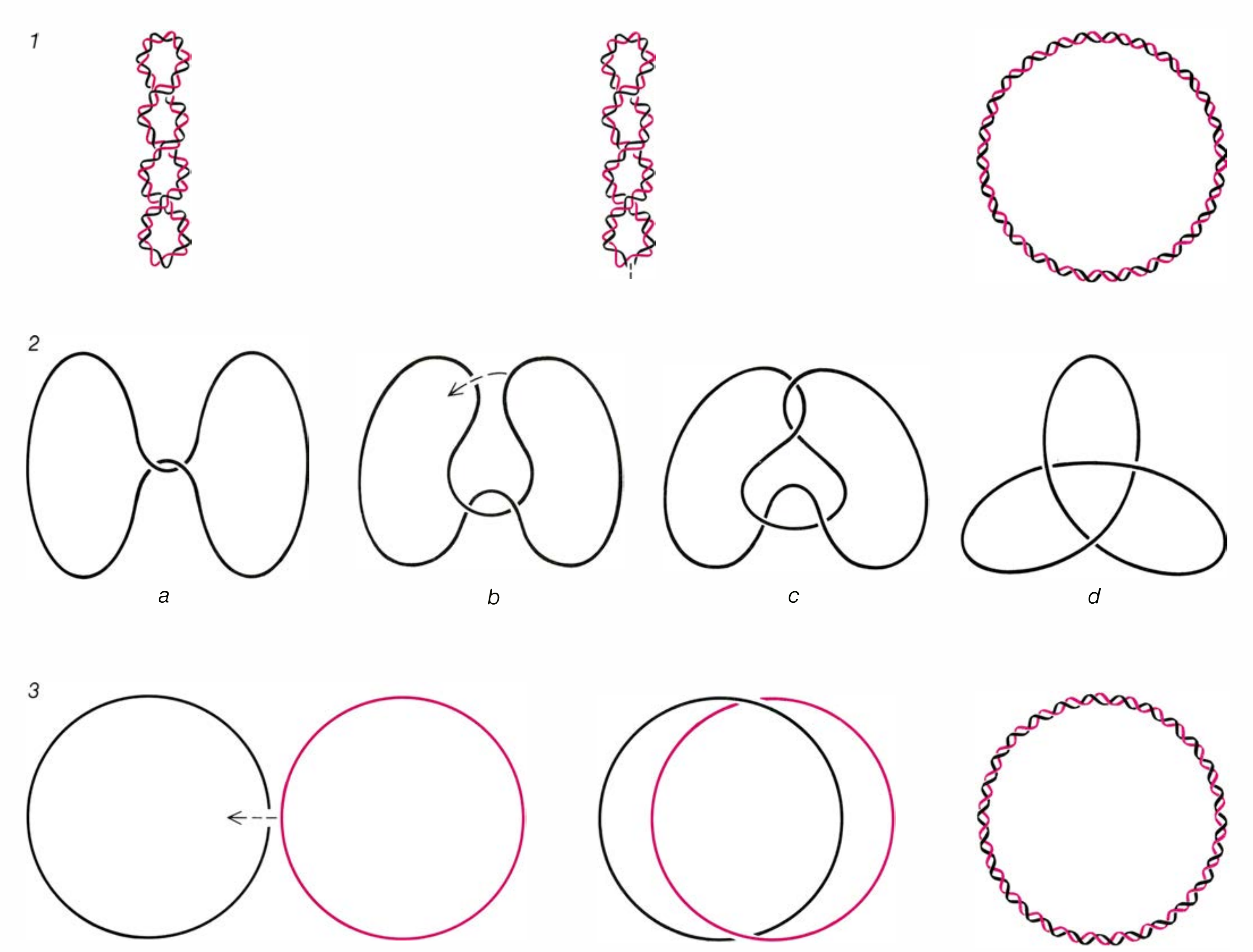
El valor del índice de enlace de un anillo se puede reducir en una proporción dada controlando la concentración del etidio. Tras la unión del etidio se emplea el enzima ligasa de ADN para reunir los cabos de la cadena cortada. La ligasa de ADN contrarresta los efectos producidos por la ADNasa. Reúne las porciones separadas de la molécula de agua que se insertó en el enlace azúcar-fosfato y elimina la molécula de agua, con lo que restablece la continuidad de la cadena. Con las dos cadenas intactas, el valor del índice de enlace queda sellado en el anillo. Cuando se retira el etidio, el anillo se superenrolla. Los anillos superenrollados aparecen al microscopio electrónico más retorcidos y más compactos que los anillos relajados. Cuanto más se haya reducido el valor del índice de enlace, más retorcido y compacto aparecerá el anillo.

El superenrollamiento que produce la reducción en el valor del índice de enlace se llama superenrollamiento negativo. Cuando el valor del índice de enlace aumenta se produce superenrollamiento positivo. Se sabe que los anillos de ADN con superenrollamiento negativo son frecuentes en las células intactas, pero los anillos con superenrollamiento positivo sólo se han preparado en el laboratorio. Conviene notar que, si se corta una cadena de un anillo superenrollado de cualquiera de los dos tipos, la molécula vuelve al estado relajado.

Aunque el superenrollamiento del ADN pueda parecer poco verosímil, existen analogías en la vida cotidiana. Si se retuercen los extremos de un tubo de goma varias veces y luego se empalman, el tubo se retuerce sobre sí mismo para formar una hélice con asas en sus dos extremos. Si se toma el auricular de un teléfono con cordón helicoidal y se le dan vueltas antes de volver a colgarlo, el cable se retuerce y forma una hélice mayor.

En el estudio de los anillos de ADN superenrollados suele utilizarse la electroforesis en gel. Cuando se colocan las moléculas de ADN en el gel y se pasa por él una corriente eléctrica, las moléculas se dirigen hacia el electrodo positivo, porque tienen una carga neta negativa. La velocidad del desplazamiento depende de la forma de la molécula, así como de su carga. Como el superenrollamiento vuelve más compacto el anillo de doble cadena, la velocidad de desplazamiento del anillo en el gel depende del valor de su índice de enlace. Después de pasar la corriente durante cierto tiempo por el gel, se puede revelar la posición de los anillos de ADN tiñendo el gel con un colorante fluorescente. Se coloca el gel sobre una fuente de radiación ultravioleta y los anillos aparecen como bandas de color rojo brillante sobre un fondo oscuro. La técnica permite distinguir fácilmente anillos que difieren en una unidad en el valor del índice de enlace.

Cuando descubrí el primer enzima capaz de llevar a cabo una transformación topológica se conocían ya desde algunos años antes el superenrollamiento del ADN y las técnicas para inducirlo



TOPOISOMERASAS DEL TIPO I, que cortan una cadena cada vez y pueden llevar a cabo varias operaciones topológicas. La primera topoisomerasa descubierta, un enzima del tipo I de *E. coli*, fue encontrada por el autor, en 1971. Sabemos ya que existen topoisomerasas en muchas especies, incluida la humana. El enzima del tipo I, al cortar una cadena en un anillo de ADN superenrollado, lleva el anillo a su estado relajado (*1*). Además, puede hacer y deshacer un nudo en la cadena simple (*2*). El nudo se forma cuando el anillo de cadena sencilla se superpone a sí mismo. Si las dos asas que así surgen se juntan, el enzima puede cortar un asa e introducir la otra por la brecha. Cuando el corte se vuelva a cerrar, el anillo formará un nudo. Las estructuras *2c* y *2d* son topológicamente equivalentes: se puede pasar de una a otra sin cortar la cadena. El enzima del tipo I también puede ensartar y desencadenar dos anillos de cadena simple (*3*). Si los anillos poseen secuencias de bases complementarias, se crea una doble hélice. Aunque todas estas operaciones parecen completamente diferentes, en cada una de ellas se requiere romper una cadena, pasar un segmento de ADN por la brecha y volver a cerrar la brecha.

en el laboratorio. Lo hallé en la conocida especie bacteriana *E. coli*. Se trata de una topoisomerasa del tipo I, perteneciente a la familia de los enzimas que cortan una sola cadena de ADN.

La primera acción de la topoisomerasa del tipo I de *E. coli* que se detectó fue la relajación de un anillo superenrollado de ADN de doble cadena. El enzima relaja con mucha mayor facilidad los anillos superenrollados negativamente que los positivos. La relajación de un anillo superenrollado constituye una de las operaciones más sencillas que puede realizar una topoisomerasa. Sólo requiere cortar una de las cadenas del anillo y volver a cerrarla. Pocos años después se descubrió que la topoisomerasa del tipo I de *E. coli* era capaz de realizar otras conversiones topológicas. Puede hacer un nudo en un anillo de cadena sencilla, y puede también deshacerlo. Además, puede catalizar el eslabonamiento de dos anillos de cadena simple complementarios para formar un anillo de cadena doble.

Leroy F. Liu ha investigado, en mi laboratorio, la formación de nudos en anillos de cadena sencilla. Sus hallazgos sirven para ilustrar las técnicas que condujeron a desvelar la acción de las topoisomerasas. Los anillos de ADN de cadena simple utilizados en los experimentos de Liu procedían del virus denominado bacteriófago *fd*, que infecta *E. coli*. Cuando se expone el ADN vírico a la topoisomerasa del tipo I de *E. coli*, en presencia de una pequeña cantidad de iones magnésicos, requisito para que actúe el enzima, se observa una transformación enigmática. Después de exponerse al enzima, el anillo de ADN de cadena simple aparece retorcido y compacto.

El trabajo experimental sugirió la formación de nudos en el anillo. El análisis por electroforesis en gel del ADN vírico sin tratar mostraba dos bandas. Una de ellas correspondía al anillo de cadena simple y la otra a la que se produce por la rotura accidental del anillo. Cuando se pone en el gel el ADN tratado con el enzima, se aprecia un grupo de bandas adicionales. Las nuevas bandas corresponden a moléculas que migran en el gel con velocidades ligeramente distintas. Pero todas las nuevas moléculas se mueven más deprisa que los anillos simples originales.

Pueden excluirse de nuestra consideración algunos hechos que explicarían estas observaciones. El medio en que

se llevó a cabo el experimento era muy alcalino; un alto grado de alcalinidad impide el emparejamiento de las bases enlazadas al esqueleto de ADN. Por tanto, las nuevas bandas del gel no pueden corresponder a estructuras creadas por el emparejamiento de bases en tramos del anillo. La microscopía electrónica revela que las nuevas estructuras son moléculas simples y por ello las diferencias de movilidad no pueden deberse a la agregación de varias moléculas de ADN.

Una hipótesis plausible es que las nuevas bandas correspondan a anillos que han formado nudos, de complejidad gradual. Existe un medio sencillo para contrastar esta hipótesis. Una vez formado el nudo en un anillo, cualquier corte producirá una cadena lineal, sin que importe la complejidad del nudo. Si las nuevas bandas creadas por tratamiento del ADN del bacteriofago *fd* con la topoisomerasa resultan de la formación de nudos, cuando se corten los anillos aparecerá una sola banda en la posición ocupada por las moléculas lineales en el gel original. Al tratar con ADNasa I el ADN de las nuevas bandas observamos, efectivamente, una banda en aquella posición.

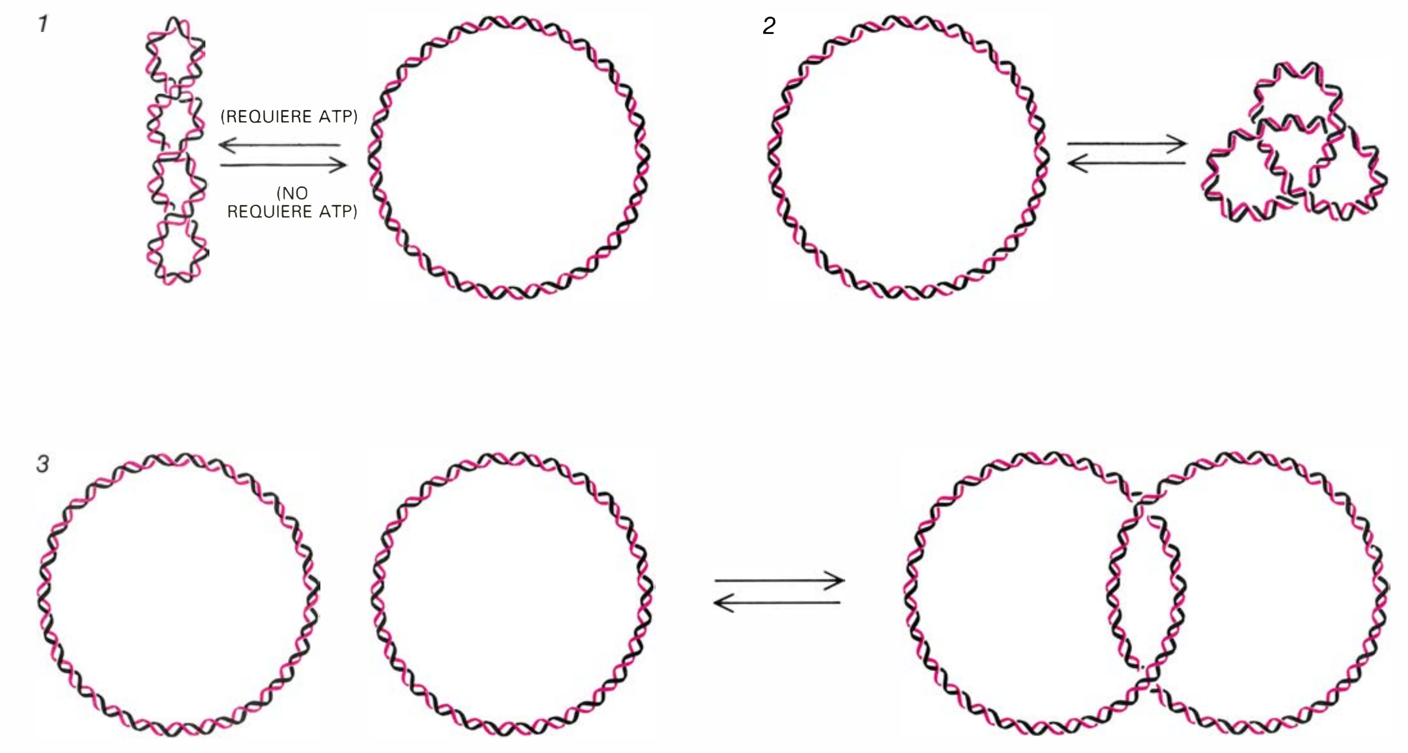
¿Cuál es el mecanismo topológico y químico a través del cual la topoisomerasa del tipo I de *E. coli* anuda un anillo de cadena simple? Parece que los segmentos complementarios del anillo monofibrilar desempeñan un importante papel en ello. Los segmentos complementarios podrían formar una doble hélice en una parte del anillo. Si se reúnen las dos asas que se proyectan de ambos extremos de la hélice, el enzima podría cortar un asa y pasar la otra por la brecha antes de cerrarla. La complejidad del nudo resultante dependería en parte de la longitud del tramo de doble hélice. Si en el anillo se hubiera formado más de una hélice, la complejidad del nudo dependería también del eslabonamiento de las diferentes asas.

Se sabe que intervienen a la vez varias moléculas de topoisomerasa sobre el mismo anillo. Pero si el anillo sufre varios cortes a un tiempo, los cabos sueltos resultantes podrían alejarse. Cabría luego una reunión incorrecta de los mismos, hasta el punto de modificar la secuencia de bases del anillo. Para averiguar si se han producido o no redistribuciones en el anillo, basta con determinar su secuencia antes y después de la actuación del enzima. Determinación que prueba que la secuencia de bases no cambia, ni siquiera cuando varias moléculas del enzima intervienen de consuno sobre el anillo. Lo que demuestra que, cuando se corta el anillo, los cabos de las cadenas de ADN no se separan a la deriva.

Es posible que el corte de una cadena ocurra siempre en aquella parte del anillo donde las bases complementarias han formado una doble hélice, la cadena intacta serviría en ese caso para mantener los cabos juntos. Y cabe también que sea el propio enzima el que mantenga los cabos juntos.

Mientras no conozcamos mejor la estructura de las topoisomerasas resultará arriesgado responder a muchos problemas que plantea el mecanismo de acción. Como ocurre con otras moléculas biológicas, las topoisomerasas presentan varios niveles de estructura. De momento sólo se ha determinado su composición química, y eso únicamente para la topoisomerasa del tipo I de *E. coli*. El enzima tiene una cadena de unos 895 residuos de aminoácido. Se ha deducido su secuencia aminoacídica a partir de la determinación de la secuencia de bases en el ADN bacteriano que codifica el enzima. Pero apenas si sabemos algo sobre la forma tridimensinal que el enzima adopta en la bacteria.

Como mencioné antes, todas las reacciones que catalizan las topoisomerasas del tipo I siguen, al parecer, el mismo curso de acción: el enzima rom-



TOPOISOMERASAS DEL TIPO II; cortan las dos cadenas de ADN a la vez y actúan solamente sobre anillos de doble cadena. El primer enzima del tipo II que se descubrió fue la llamada girasa de ADN. Se encontró en *E. coli*. En presencia de ATP, puede superenrollar negativamente un anillo de doble cadena relajado (*1*). Sin ATP puede relajar un anillo superenrollado negativamente, pero la reacción progresa lentamente. El enzima del tipo II puede hacer y deshacer nudos en un anillo de cadena doble, como hace el enzima del tipo I con un anillo de cadena simple (*2*). Sobre un mismo anillo pueden trabajar a la vez muchas moléculas de topoisomerasa. Si los cabos se pudieran separar a la deriva durante la formación de nudos, la secuencia de bases del anillo podría redistribuirse cuando se volvieran a unir los extremos. No se han detectado tales redistribuciones. Cuando se corta el anillo, el enzima forma un enlace covalente con la cadena de ADN. Este enlace impide que los cabos se separen. Es capaz, también, de desencadenar dos anillos de doble cadena.

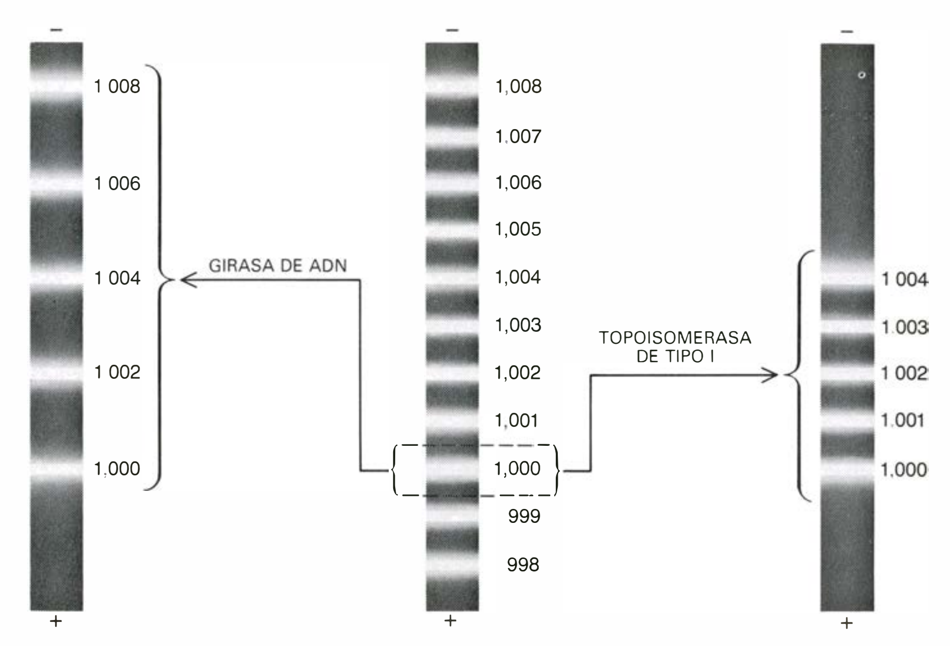
pe el esqueleto del ADN, introduce un segmento de ADN por la brecha abierta y, finalmente, el enzima cierra la brecha que él mismo produjo. Hay varias maneras de acometer esas tres operaciones. Una sería que la topoisomerasa incorporara en una misma molécula las actividades de los enzimas ADNasa I y ligasa. La topoisomerasa rompería el enlace entre el grupo azúcar y el grupo fosfato, mediante la inserción de una molécula de agua, y luego restablecería el puente eliminando la molécula de agua.

Sin embargo, la energética de los mecanismos combinados de la ADNasa y la ligasa sugiere que no es ese el camino que sigue la topoisomerasa. La formación y rotura de enlaces químicos pueden absorber o liberar energía. Una reacción dada puede progresar en uno u otro sentido según sea el balance de energía absorbida o liberada. Los enzimas no determinan la dirección de una reacción; sólo la aceleran en el sentido en que liberan energía.

La rotura del enlace azúcar-fosfato por la ADNasa diríamos que goza del favor energético. La reacción puede ocurrir de manera espontánea. La ADNasa se limita a acelerarla. Por el contrario, la regeneración del enlace azúcar-fosfato va acompañada de un balance energético desfavorable; no ocurre además de manera espontánea. La ligasa de ADN sólo restituye el enlace gracias a una reacción auxiliar con una molécula que se denomina cofactor de alta energía. La escisión del cofactor proporciona la energía que se necesita para volver a cerrar el enlace azúcar-fosfato. Las ligasas de ADN bacterianas emplean como cofactor el dinucleótido de niacinamida y adenina (NAD). Las ligasas de otros organismos utilizan adenosín trifosfato (ATP).

Las topoisomerasas del tipo I prescinden del cofactor para actuar, por lo que es poco probable que el esqueleto del ADN se refunda a través de un mecanismo como el de la ligasa. Ante la ausencia de cofactor pensé que las topoisomerasas del tipo I cortan y reajustan las cadenas de ADN a través de un mecanismo totalmente distinto. En vez de añadir porciones de la molécula de agua a los grupos de azúcar y de fosfato, podría escindirse la propia topoisomerasa: una pequeña porción del enzima, verbigracia, un átomo de hidrógeno, se uniría al azúcar, y el resto del enzima, por enlace covalente, al grupo fosfato.

Si este fuera el mecanismo empleado para las topoisomerasas del tipo I, la rotura transitoria de la cadena del ADN debería ir acompañada de la formación de un complejo covalente de ADN y enzima. Todos los esfuerzos encaminados al aislamiento de tal complejo resultaron infructuosos durante años. Pero se acaba de comprobar que al añadir a una mezcla de ADN y topoisomerasa una sustancia que deshace el plegamiento de las proteínas, suelen aparecer tales complejos. Se ha logrado establecer la naturaleza del enlace entre el enzima y la cadena de ADN en tres topoisomerasas bacterianas. En todas ellas aparece un grupo fosfato del extremo 5' de la cadena unido por medio de un enlace covalente a un residuo de tirosina del enzima.



LA GIRASA DE ADN rebaja el valor del índice de enlace de un anillo de doble cadena en dos unidades o múltiplos de dos. Se pueden separar las moléculas de ADN en razón del valor de su índice de enlace por medio del método de la electroforesis en gel, en el que las moléculas migran, a través de un gel, hacia el electrodo positivo. Cuanto más difiere el valor del índice de enlace del que corresponde al estado relajado, tanto más rápidamente migra el anillo. La razón estriba en que cuanto más superenrollado se encuentre un anillo, tanto más retorcido y compacto será. Por un procedimiento fotográfico se logra que las moléculas aparezcan en forma de bandas claras sobre un fondo oscuro. La posición de la banda señala el valor del índice de enlace del anillo. Se prepara una familia de topoisómeros con distintos índices de enlace mezclando el ADN con etidio. Si se hace correr la familia de topoisómeros en un gel, forma un grupo de bandas estrechamente espaciadas (*centro*). Cortando una banda del gel se aisla un conjunto de anillos con el mismo valor del índice de enlace, por ejemplo 1000. Si el corte del gel se trata con topoisomerasa del tipo I de *E. coli*, el enzima relaja los anillos elevando el valor del índice de enlace de unidad en unidad (*derecha*). Sin embargo, si el corte se trata con girasa de ADN en ausencia de ATP, aparecen bandas que corresponden a anillos cuyo valor del índice de enlace ha aumentado en un número par de unidades.

Se tardó varios años, desde el descubrimiento de las topoisomerasas del tipo I de *E. coli*, en hallar la primera topoisomerasa del tipo II. Lo logró, en 1976, Martin Gellert, del Instituto Nacional de la Salud de los Estados Unidos. El descubrimiento se produjo en un estudio sobre la infección vírica de una célula bacteriana. El ADN del virus bacteriófago λ, que infecta *E. coli*, puede formar un anillo de doble cadena después de penetrar en la bacteria. Bajo ciertas condiciones, el ADN vírico se inserta en el ADN del huésped. El enzima denominado integrasa λ, codificado por el ADN del virus, cataliza esta última reacción. La integrasa corta el anillo vírico de ADN y el ADN de la bacteria; suelda luego los cabos, de modo que el ADN del virus se convierte en un segmento del cromosoma bacteriano.

En 1975, Howard A. Nash, del Instituto Nacional de Salud Mental, y Kiyoshi Mizuuchi, del Instituto Nacional de la Salud, lograron en el laboratorio esa inserción con integrasa λ purificada. Observaron que si el anillo del virus estaba inicialmente relajado se necesitaba ATP para la inserción; no se requería, en cambio, ATP cuando el anillo se encontraba superenrollado negativamente. Nash, Mizuuchi y sus colegas Geller y Mary H. O'Dea aventuraron que el anillo debía estar superenrollado para que tuviera lugar la inserción, y que el enzima que llevaba a cabo el superenrollamiento necesitaba ATP como cofactor. Su hipótesis se ha confirmado.

En 1976, Gellert y sus colegas aislaron de *E. coli* el enzima superenrollante. Se le denominó girasa de ADN, y puede conferir una serie de formas topológicas a los anillos de doble cadena. Las girasas bacterianas son, que se sepa, las únicas topoisomerasas capaces de superenrollar un anillo relajado de ADN. Para el superenrollamiento se

necesitaba ATP como cofactor. La rotura del ATP por medio de la inserción y escisión de una molécula de agua, que libera energía, proporciona la fuerza motriz que necesita el proceso de superenrollamiento, que consume energía. En ausencia de ATP, la girasa de ADN es capaz de relajar un anillo superenrollado negativamente, si bien con gran lentitud.

Cuando se descubrió la girasa, se pensó que cortaba, en la doble hélice, una cadena cada vez, como hacían todas la topoisomerasas conocidas hasta entonces. En 1977, el grupo de Gellert y el de Nicholas R. Cozzarelli, de la Universidad de Chicago, trataron disoluciones de ADN y girasa con una sustancia que perturbaba el plegamiento de las proteínas. Al aplicar la sustancia, los anillos adoptaban una forma lineal. El enzima rompía ambas cadenas de ADN a la vez. Pudo establecerse incluso que una parte de la molécula del enzima se unía por medio de un enlace covalente al extremo libre 5' de cada una de las dos cadenas.

Mas no se apreció en todo su significado la escisión de ambas cadenas hasta que se descubrieron otras dos reacciones que catalizaba también la girasa. En 1980, Leroy Liu, Chung-Cheng Liu y Bruce M. Alberts, de la Universidad de California en San Francisco, descubrieron que una topoisomerasa codificada por el ADN del virus bacteriano T4 podía hacer nudos en un anillo de doble cadena. Casi al mismo tiempo, Cozzarelli y Kenneth N. Kreuzer, de la Universidad de Chicago, probaron que la girasa podía encadenar dos anillos de doble cadena y también separar dos anillos encadenados.

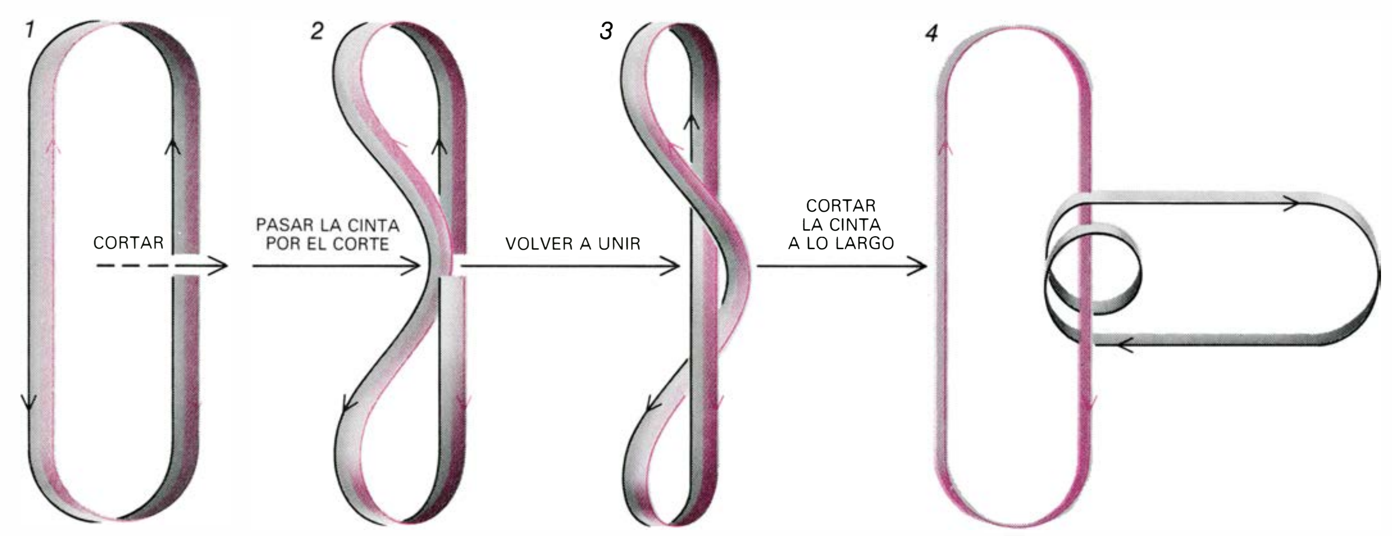
Para deshacer un nudo en un anillo de doble cadena es preciso cortar ambas cadenas y pasar un segmento del ADN por la brecha. Para separar anillos encadenados también deben cortarse las dos cadenas del anillo. Sin embargo, cabría que el enzima no cortase las dos cadenas a la vez. La girasa podría cortar una cadena, pasar el segmento a la médula de la doble hélice y reunir los cabos, soldándolos antes de cortar la segunda cadena. Pero un mecanismo así parece demasiado rebuscado en comparación con un corte transitorio de ambas cadenas.

La argumentación más convincente de que la girasa corta las dos cadenas a la vez deriva del examen de la forma en que el enzima altera el valor del índice de enlace. Cuando se exponen a topoisomerasa del tipo I anillos superenrollados negativamente con un índice de enlace de 1000, el valor del índice aumenta gradualmente, en pasos de una unidad, hasta alcanzar el que corresponde al estado relajado. Si la reacción se interrumpe antes de que se alcance el estado relajado, la electroforesis nos revela una familia de topoisómeros con distintos valores del índice de enlace, todos superiores a 1000. Vinograd y sus colegas fueron los primeros en llevar a cabo un experimento de este tipo, en 1975. Encontraron que los anillos tenían valores del índice de enlace que se apartaban del valor original en números enteros y que podían diferir entre sí hasta sólo en una unidad.

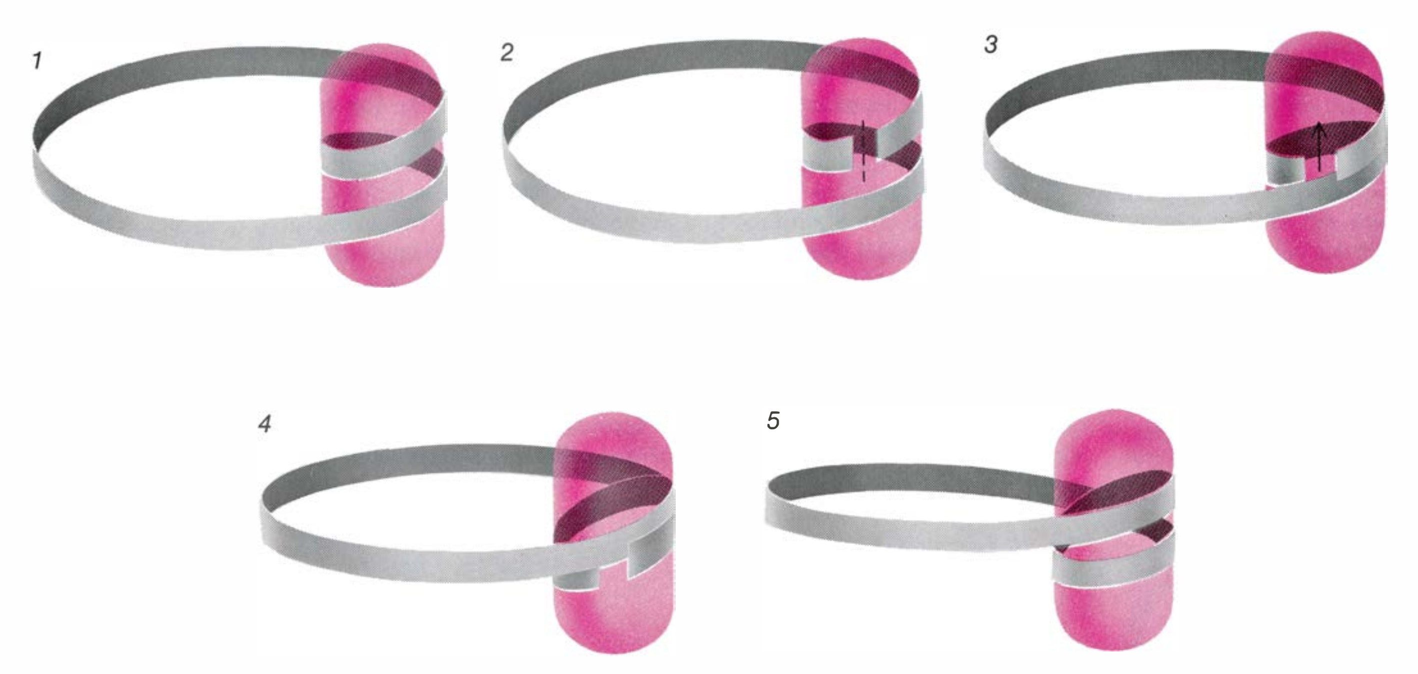
No obstante, al exponer anillos de doble cadena a una topoisomerasa del tipo II, sólo se observaron topoisómeros con valores de índice de enlace que superaban al valor del índice de enlace de los anillos originales en dos o múltiplos de dos unidades. Si la reacción se realizaba en presencia de ATP, el enzima superenrollaba el anillo, en lugar de relajarlo, pero otra vez la variación en el valor del índice de enlace ocurría en pasos de dos unidades.

Un ejercicio sencillo ilustrará cómo consigue la girasa reducir el valor del índice de enlace de un anillo de doble cadena en dos unidades o múltiplos de dos. Tomemos un fragmento de cinta. Si unimos los extremos para formar un anillo, los bordes de la cinta adoptan la forma de asas circulares cerradas. Los bordes pueden representar las cadenas complementarias del ADN, una que marcha en el sentido 5' a 3' y la otra que lo hace en el sentido de 3' a 5'. Si los extremos de la cinta no se retuercen antes de unirlos, el índice de enlace de las dos asas que describen los bordes valdrá cero. Se puede comprobar fácilmente que el valor es cero cortando en dos la cinta a lo largo de su eje longitudinal. Las dos mitades, cada una de las cuales contiene uno de los bordes originales, se separarán.

Supongamos ahora que, en lugar de cortar la cinta longitudinalmente, se haga un corte transversal. Los cabos se sujetan manteniéndolos cerca, sin girar. Luego, se pasa por la brecha abier-



MODELO DE CINTA que ilustra de qué manera la girasa modifica el valor del índice de enlace de un anillo de doble cadena en múltiplos de dos unidades. Los extremos de la cinta se reúnen, sin girarlos, para formar un solo bucle. Los bordes del bucle o asa representan las cadenas de ADN que marchan en sentidos opuestos. Una va de 5' a 3' y, la otra, de 3' a 5'. Puesto que los bordes no se cruzan, el valor del índice de enlace es cero. Se corta la cinta transversalmente (*1*) y se sujetan los cabos sin girarlos. Se pasa un segmento del asa por la brecha (*2*) y se sueldan los cabos (*3*). Si cortamos entonces la cinta longitudinalmente, apreciaremos que un borde da dos vueltas completas alrededor del otro. El valor del índice de enlace ha aumentado, pues, de cero a dos (*4*). Se puede probar que la operación altera siempre en dos unidades el valor del índice de enlace. Que el valor aumente o disminuya dependerá en todo caso de la forma exacta en que se pase el segmento por la brecha. Pero si se gira 360 grados uno de los cabos de la cinta antes de soldarla, el valor del índice de enlace variará en una unidad, porque se habrá añadido una revolución. Puesto que, según se ha observado, la girasa de ADN cambia siempre el valor del índice de enlace en un número par de unidades, el enzima tiene que ingeniárselas de alguna manera para impdeir la rotación de los extremos.



MECANISMO DE ACCION propuesto para la girasa de ADN. Un anillo de ADN envuelve en principio la molécula de enzima con una gran espira dextrógira (*1*). El enzima se une a un segmento de ADN de unos 150 pares de bases y corta el anillo, rompiendo el enlace covalente que une un grupo de azúcar y un grupo de fosfato en cada una de las cadenas del esqueleto del ADN (*2*). A continuación se pasa por la brecha un segmento del anillo (*3, 4*). Al cerrar la brecha, el valor del índice de enlace ha variado en dos unidades (*5*). Por añadidura, el paso por la hendidura ha invertido el sentido de giro de la gran espira de ADN. Se cree que, mediante un proceso desconocido, la espira de ADN torna al sentido que tenía originalmente antes de reiniciar la reacción.

ta un tramo opuesto de la cinta y los cabos se pegan sin girar. En esta operación, el valor del índice de enlace ha variado en dos unidades, pasando de cero a dos. Esto es fácil de comprobar cortando la cinta a lo largo de su eje longitudinal. Se obtienen entonces dos anillos, uno de los cuales rodea al otro por dos veces [*véase la ilustración de la página anterior*]. Este ensayo constituye un caso particular, ya que el valor original del índice de enlace era cero. Pero puede probarse que la variación en el valor del índice de enlace es siempre dos, sin que importe el valor inicial del índice de enlace. Este método para producir variaciones pares en el valor del índice de enlace lo señaló ya el matemático Brock Fuller, del Instituto de Tecnología de California, antes de que se realizara el trabajo sobre las topoisomerasas del tipo II.

Los experimentos efectuados con la girasa de ADN tienden a confirmar la idea de que el enzima altera el valor del índice de enlace del anillo de doble cadena por un mecanismo semejante, en líneas generales, al que ilustra el ejemplo de la cinta cortada. Conviene subrayar que una de las condiciones del ejemplo es que los extremos cortados no se giren. Si un extremo girase 360 grados alrededor del eje longitudinal de la cinta, el valor del índice de enlace variaría en una unidad: un número impar. Por tanto, la girasa de ADN debe impedir de algún modo que los cabos giren mientras transcurre la reacción.

El trabajo realizado en mi laboratorio y en otros ha probado que la girasa de ADN es una proteína con un peso molecular aproximado de 400.000. La proteína contiene dos pares idénticos de subunidades, designados *A* y *B*. Si se desnaturaliza el enzima mientras está cortando el anillo, las subunidades *A* aparecen unidas a los extremos libres 5' de las cadenas. Este resultado sugiere el motivo por el que los cabos del anillo no se separan durante la rotura de las cadenas. Los extremos 5' de las cadenas permanecen enlazados a las subunidades *A*. A su vez, las subunidades *A* permanecen juntas gracias a las subunidades *B*. Otros trabajos adicionales han probado que el enzima se une a un tramo de ADN de 150 pares de bases. La mayor parte de ese tramo de ADN parece envolver la molécula del enzima.

Se han propuesto varios mecanismos para explicar cómo altera la girasa el valor del índice de enlace. Cozzarelli y Patrick O'Brown sugieren que un segmento del anillo de ADN forma una espira dextrógira que envuelve la molécula del enzima. La unión de una molécula de ATP al complejo que forman el anillo de ADN y el enzima determina que una porción del anillo atraviese una brecha abierta en la espira de ADN situada sobre la superficie de la molécula de la girasa. A continuación se cierra la brecha. De ese modo, el paso a través de la brecha no sólo cambia en dos unidades el valor del índice de enlace, sino que invierte también el sentido de la espira de ADN que envuelve la molécula del enzima. [*Véase la ilustración de esta misma página.*] Después de la escisión de la molécula del ATP, la espira de ADN retornaría a su forma dextrógira original.

En un segundo mecanismo propuesto, la molécula de ADN configura también una espira con la que envuelve la molécula del enzima, pero el sentido de esta espira no cambia durante la reacción. Aquí, un segmento de ADN se aproxima al que yace sobre la superficie del enzima. La rotura del anillo de doble cadena ocurre cerca del centro del segmento que yace sobre el enzima. El segmento intacto atraviesa entonces la brecha. La unión con una molécula de ATP cambia la forma del enzima, para permitir que el segmento intacto de ADN pase al interior de la molécula del enzima. La escisión de la molécula de ATP hace que el segmento intacto de ADN salga por el otro lado del enzima [*véase la ilustración superior de la página siguiente*]. Cualquiera de estos dos mecanismos puede explicar las transformaciones topológicas que provoca la girasa. Resta mucho trabajo por hacer antes de determinar cuál de ellos sea el correcto.

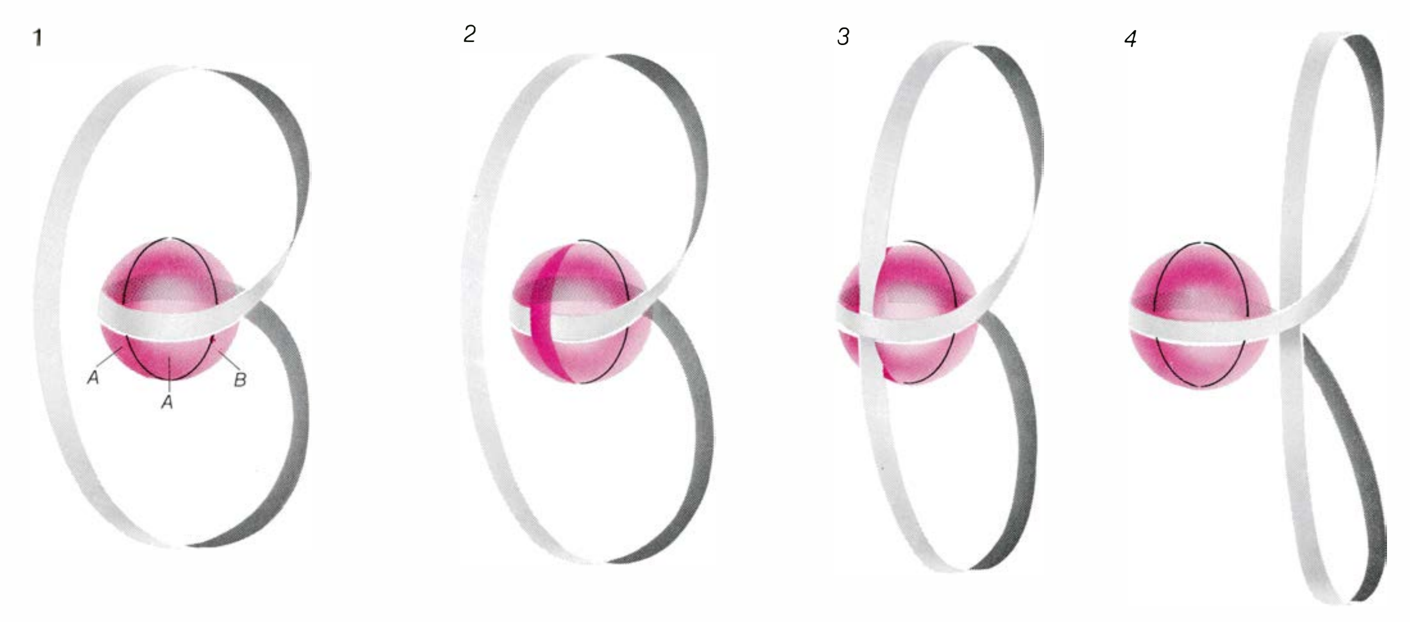
Mientras se investigaba la acción de las topoisomerasas a nivel molecular se encontraron estos enzimas en muchas especies, incluida la humana. Parece que todos los organismos, procariotas y eucariotas, poseen topoisomerasas del tipo I y del tipo II. (Los procariotas son organismos, como las bacterias, cuyo material genético no es-

tá encerrado en un núcleo. Las células de los organismos eucariotas poseen núcleo diferenciado.)

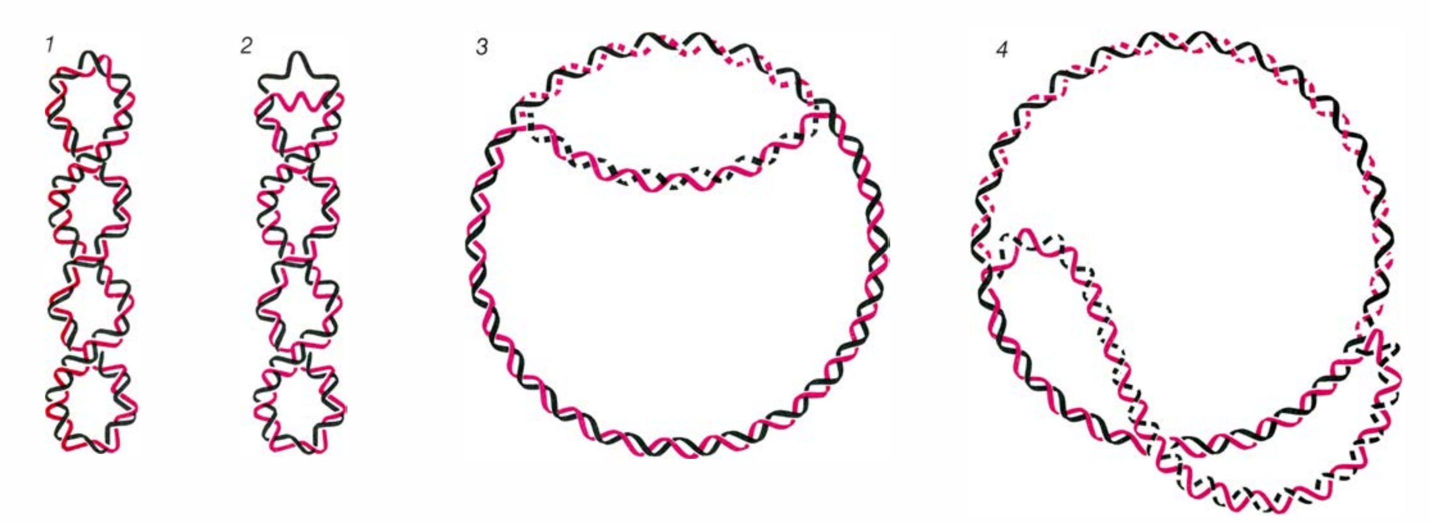
Las topoisomerasas de los eucariotas difieren en cierta medida de las que presentan los procariotas. Poco después del descubrimiento de la topoisomerasa del tipo I de *E. coli*, James J. Champoux y Renato Dulbecco, del Instituto Salk de Estudios Biológicos, hallaron un enzima del tipo I en el núcleo de células de ratón. La topoisomerasa de ratón relaja con facilidad anillos de ADN de doble cadena superenrollados positiva o negativamente, en contraste con el enzima bacteriano, que relaja con más facilidad los anillos superenrollados negativamente. Aún más, cuando el enzima de ratón corta el esqueleto de ADN forma un enlace covalente con el extremo 3' de la cadena del ADN y no con el extremo 5'. Las topoisomerasas del tipo I de otros eucariotas parecen compartir estas propiedades.

También se han encontrado en muchas especies topoisomerasas del tipo II. Pero sólo el enzima bacteriano girasa de ADN puede superenrollar anillos de ADN. Por otro lado, las topoisomerasas del tipo II de virus y eucariotas llevan a cabo las demás operaciones que realiza la girasa de ADN: formación de nudos y eslabonamiento de anillos de doble cadena.

La amplia distribución de que gozan las topoisomerasas lleva naturalmente a preguntarse sobre cuál sea su función biológica. La información disponible para responder a esa cuestión procede sobre todo de las investigaciones realizadas con bacterias y virus que las infectan. Aunque limitados, los datos con que se cuenta sugieren que los enzimas intervienen en la replicación del ADN,



UNA COMPUERTA MOLECULAR desempeña un importante papel en un segundo mecanismo propuesto para la girasa de ADN. Sabemos que la molécula de la girasa está constituida por dos pares de subunidades idénticas, designadas *A* y *B*. Un segmento de unos 150 pares de bases envuelve al enzima en una gran espira dextrógira (*1*). Las dos subunidades *A* constituyen la superficie sobre la que yace la parte central del segmento de 150 pares de bases. Las dos subunidades *A* tiran del anillo y lo rompen (*2*). El segmento vertical del anillo pasa por entre las dos subunidades *A*, que actúan de compuerta, y llega al interior del enzima (*3*). Se suelda entonces la brecha de la hélice. Cuando el segmento vertical ha recorrido la mitad del enzima, la compuerta se cierra, momento en que las subunidades *B* se abren. Actúan así de compuerta de salida por donde se expulsa del enzima el segmento vertical (*4*).



LA REPLICACION de la información contenida en la secuencia de pares de bases es imprescindible para la división de la célula. La replicación de un anillo de doble hélice de ADN quizá dependa de la capacidad de una topoisomerasa del tipo II para reducir el valor del índice de enlace. Un mecanismo común en la duplicación de un anillo de doble cadena es la formación de una cadena complementaria de cada una de las cadenas originales. Para que se ensamblen las nuevas cadenas deben separarse las originales. En las bacterias, la girasa superenrolla negativamente el anillo (*1*), lo que facilita la apertura de la doble hélice. Es probable que en algunos organismos este desenrollamiento sea necesario para la replicación (*2*). El montaje de las nuevas cadenas comienza utilizando como modelo las cadenas originales (*3*). Para que avance la replicación, debe abrirse progresivamente la doble hélice. En este proceso quizá actúen una o más moléculas de topoisomerasa, que reducirían repetidamente el valor del índice de enlace (*4*). Cuando el valor del índice de enlace es cero, los anillos de la doble cadena se separan libremente. Cada anillo queda de esta forma constituido por una cadena original y otra nueva.

en la transcripción de la información genética y en ciertos tipos de recombinación de los genes.

Diríase que el ADN de la mayoría de las especies bacterianas está superenrollado negativamente, por lo general. El grado de superenrollamiento viene controlado por las topoisomerasas. La girasa superenrolla el ADN; una topoisomerasa del tipo I puede relajarlo en la proporción requerida. La acción equilibrada de los dos enzimas puede tener implicaciones significativas. Tanto para la replicación de la información genética como para su transcripción en forma de ARN deben separarse las cadenas de la doble hélice mediante la rotura de los enlaces de hidrógeno que unen los pares de bases. La separación expone las bases de cada cadena a la acción de los enzimas. El superenrollamiento negativo facilita el desenrollamiento, y la separación por tanto, de las cadenas del anillo. La razón estriba en que un modo por medio del cual la doble hélice se puede acercar al estado relajado es desenrollándose.

Cuando parte de la doble hélice se deshebra, dejando libre un segmento en cada cadena, el valor del índice de enlace del anillo permanece inalterado, ya que las cadenas permanecen intactas. El desenrollamiento, sin embargo, vuelve a distribuir las revoluciones que dan las dos cadenas una alrededor de la otra. Las revoluciones quedan confinadas esencialmente a la parte del anillo que conserva la configuración helicoidal. Puesto que el mismo número de revoluciones se distribuyen en una longitud menor, la relación del número de revoluciones al número de pares de bases puede acercarse, en el segmento helicoidal, al valor del anillo relajado: una por cada 10,5 pares de bases.

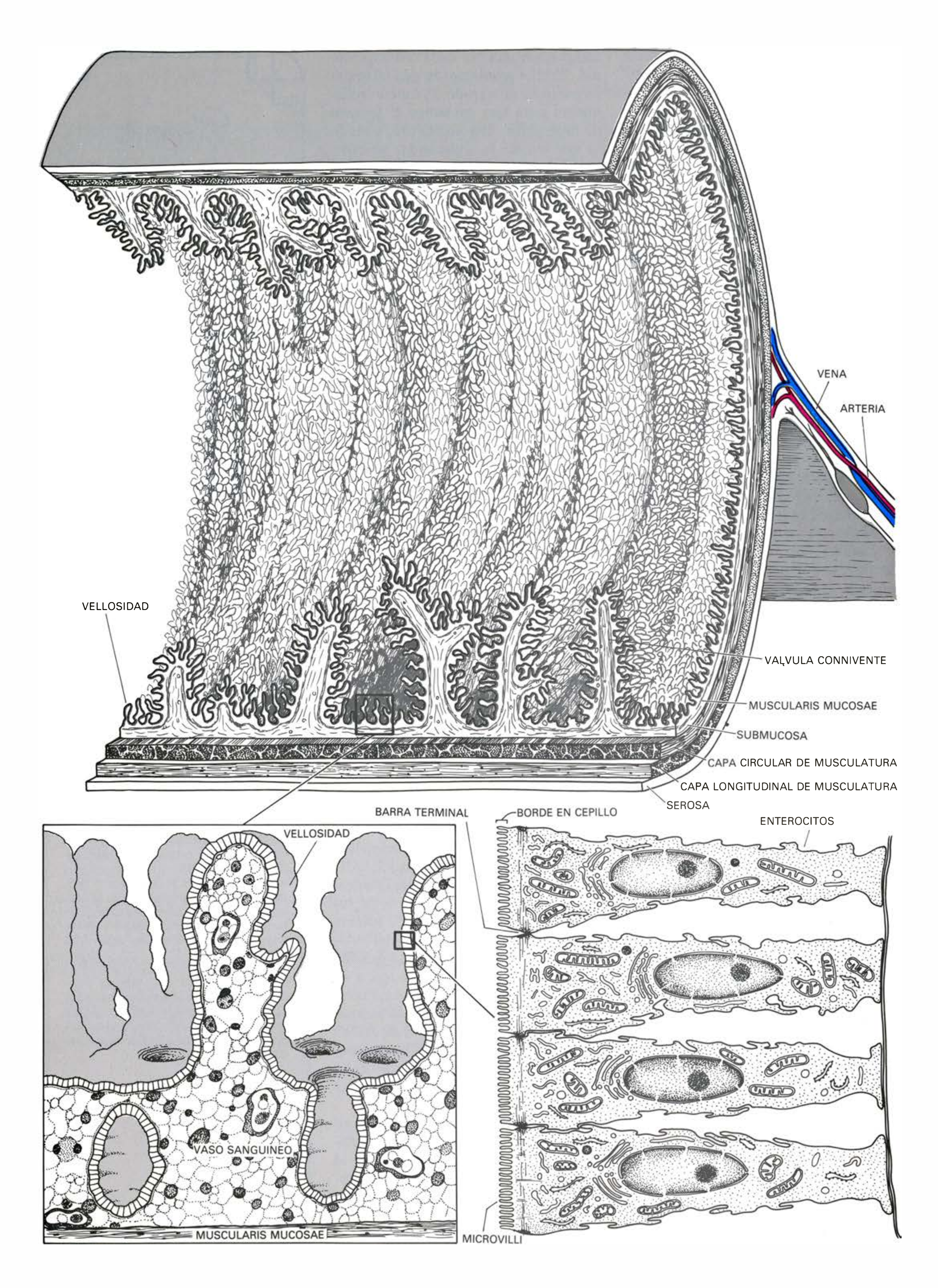
Al facilitar la separación de las dos cadenas de la doble hélice, el superenrollamiento favorece la replicación y la transcripción. Por ejemplo, un enzima codificado por un gen denominado *A* del anillo de ADN del virus ϕX174 desempeña un importante papel en el ciclo vital del virus. Mediante la rotura del anillo de ADN, el enzima inicia la replicación de la partícula del virus en la célula huésped. En el laboratorio, el enzima del gen *A* corta el ADN vírico sólo si el ADN está superenrollado negativamente; puede que *in vivo* también ello sea necesario.

La replicación de la información genética previa a la división celular, tanto en bacterias como en células eucariotas, necesita asimismo que se separen las cadenas. A medida que lo van haciendo, se forma una nueva cadena complementaria por cada cadena original. Habrá organismos en que tal vez se requiera un alto grado de superenrollamiento para que comience el proceso de replicación. Por añadidura, a medida que avanza la replicación, las cadenas del ADN deben progresivamente reducir su torsión, disminuyendo así el valor del índice de enlace [*véase la ilustración inferior de la página anterior*]. Para que los nuevos anillos de doble cadena se puedan separar al final de la replicación, el valor del índice de enlace ha de ser nulo. Si el valor no es exactamente cero, los anillos permanecerán eslabonados. Tal eslabonamiento se ha observado en el ADN vírico al final del proceso de replicación; quizá la etapa última, en que quedan sueltos, la lleve a cabo una topoisomerasa.

De manera análoga, la función de las topoisomerasas en la transcripción exigirá, presumiblemente, la actuación de su capacidad para abrir la doble hélice. La transcripción se inicia con la unión del enzima ARN polimerasa a la molécula de ADN. El enzima separa unos 10 pares de bases cada vez para permitir que se transcriba la secuencia de bases. El superenrollamiento facilita la separación de las cadenas. La unión de la ARN polimerasa a secuencias del ADN se ve afectada por la unión de otras proteínas en centros adyacentes; en muchos casos la unión de esas otras proteínas depende del grado de superenrollamiento. Así, el ajuste del grado de superenrollamiento contribuiría a la regulación de la tasa de transcripción, que es un elemento crucial del metabolismo celular.

La rotura y reunión de las cadenas de ADN también constituye un paso esencial de la recombinación y la redistribución génicas. La integración del ADN vírico en el ADN del huésped, catalizada por la integrasa λ, es uno de los ejemplos mejor estudiados. El enzima corta el ADN vírico y el ADN del huésped; los cabos del ADN vírico se unen a los cabos del ADN del huésped, y viceversa. En este sentido, Nash y Yoshiko Kikuchi han mostrado que la integrasa λ es una topoisomerasa de ADN: el enzima puede relajar ADN superenrollado negativamente y forma un enlace covalente con el ADN cuando corta las cadenas.

Cabe esperar que procesos biológicos tan complicados como los que llevan a cabo las topoisomerasas sobre el ADN resulten ser funciones de vital importancia. A medida que avancen los estudios sobre las topoisomerasas se demostrará, sin duda, que su implicación en los procesos genéticos de la célula es todavía más profunda.



VENA
ARTERIA
VELLOSIDAD
VALVULA CONNIVENTE
MUSCULARIS MUCOSAE
SUBMUCOSA
CAPA CIRCULAR DE MUSCULATURA
CAPA LONGITUDINAL DE MUSCULATURA
SEROSA
BARRA TERMINAL
BORDE EN CEPILLO
ENTEROCITOS
VELLOSIDAD
VASO SANGUINEO
MUSCULARIS MUCOSAE
MICROVILLI

La secreción intestinal

Un mismo tipo celular es capaz de absorber y segregar agua y electrolitos. El balance entre los dos procesos depende de la concentración de calcio en la célula

Anunciación Ilundáin y R. J. Naftalin

Los epitelios son tejidos formados por una o varias capas de células, dispuestas en láminas, que cubren superficies o tapizan cavidades del organismo. Permiten, además, que las soluciones de un lado y otro de la pared sobre la que se apoya el epitelio mantengan una composición diferente. Desempeñan, pues, un papel crucial en la homeostasis del organismo. (Se llama homeostasis a la conservación fisiológica de las propiedades químicas y físicas del medio interno de un ser vivo en un estado constante.) Como parte de ese mecanismo homeostático, la capacidad de transporte de los epitelios está cuidadosamente regulada por factores hormonales y nerviosos.

La pared del intestino delgado consta de capas sucesivas de tejido muscular, tejido conectivo y epitelio simple, que es el que controla el tránsito de materiales a través de dicha barrera intestinal. Los enterocitos, o células epiteliales responsables del trasiego selectivo de sustancias a través de la pared intestinal, ofrecen al microscopio electrónico una estructura asimétrica. Son células altas, columnares, cuya superficie apical (que mira hacia la luz del intestino) presenta, a manera de proyecciones digitales, un gran número de microvellosidades, por lo que recibe el nombre de borde en cepillo.

Las membranas que ciñen a los enterocitos dejan espacios de separación intercelulares, si bien presentan en el extremo luminal complejos de unión, que separan el espacio intercelular de la luz del intestino, al tiempo que contribuyen a mantener la estructura del epitelio. La distribución de iones a un lado y otro de la membrana que rodea al enterocito no es tampoco simétrica. La concentración de sodio (Na^+) y cloruro (Cl^-) es más alta en el medio extracelular que en el intracelular; al revés de lo que ocurre con el potasio (K^+). Existe, además, entre el lado serosal o sanguíneo y el lado mucosal o luminal del epitelio, una diferencia de potencial eléctrico. También hay diferencia de potencial entre el interior de la célula y el exterior, a través del borde en cepillo y de la membrana basolateral.

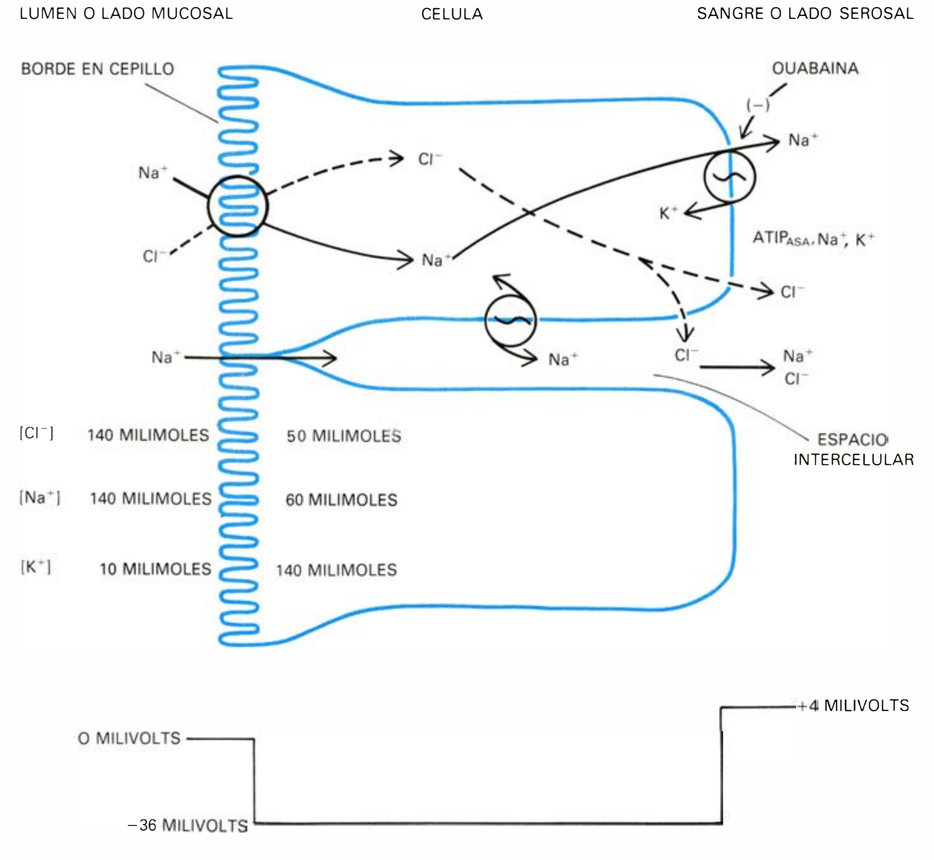
Tenemos, pues, unos gradientes iónicos y unos gradientes eléctricos. ¿Cómo se mantienen? Fundamentalmente, gracias a la actividad de una bomba localizada en la membrana basolateral, que impulsa sodio hacia el medio extracelular y potasio hacia el intracelular, en contra de sus gradientes electroquímicos. La bomba en cuestión es un enzima: un ATPasa dependiente de Na^+ y K^+ (la ATPasa Na^+ y K^+). Los glucósidos cardíacos, la ouabaina por ejemplo, inhiben la acción catalítica del enzima.

La cantidad neta de solutos y agua que atraviesa la pared del intestino es el resultado de dos movimientos, uno de entrada hacia la sangre y otro de salida hacia la luz del intestino. Si una sustancia pasa de la luz intestinal a la sangre en mayor cantidad que de ésta al lumen, se absorbe la sustancia. Cuando ocurre el caso inverso, tenemos secreción.

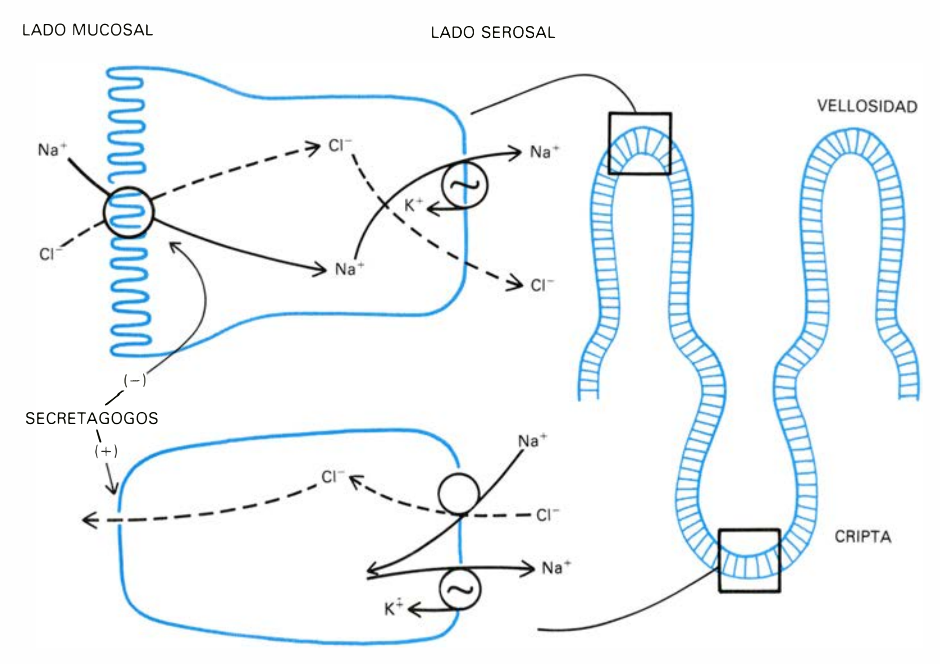
El epitelio intestinal posee la rara habilidad de absorber y secretar simultáneamente, aunque suele dominar la absorción con un paso neto a la sangre de sustancias orgánicas, iones y agua. El Na^+ penetra en el interior celular a través del borde en cepillo, a favor del gradiente de potencial electroquímico (transporte pasivo); sale luego contra gradiente (transporte activo) al lado serosal y espacios intercelulares bombeado por la ATPasa Na^+ y K^+. La absorción intestinal de muchos solutos y agua está pasivamente ligada al transporte activo del sodio; la solución absorbida es isotónica con el plasma de los capilares sanguíneos (ambos fluidos poseen la misma presión osmótica). En numerosos epitelios se ha demostrado que la entrada de Cl^- en las células por su membrana apical está acoplada, en forma eléctricamente neutra, al movimiento del sodio. Se cree que, debido al gradiente de Na^+, el cloruro se acumula en la célula por encima de su potencial de equilibrio electroquímico y sale al lado serosal y espacio intercelular a favor de gradiente. Las uniones intercelulares descritas antes son permeables a los cationes (iones dotados de carga positiva), pero muy poco a los aniones (iones dotados de carga negativa), lo que explica que el paso del Na^+ a través de esas uniones sea más importante que el del Cl^-. El movimiento de agua a través de la pared intestinal acompaña al de iones.

La capacidad secretora del intestino delgado se describió ya a principios del siglo XX. Pero hubo que esperar hasta las postrimerías de la década de los sesenta para que los fisiólogos cayeran en la cuenta de su interés. Lo que tuvo que ver con el deseo de explicar la voluminosa diarrea que se produce en los pacientes afectados de cólera. Cuando el microorganismo *Vibrio cholerae* –inductor del cólera– coloniza el intestino, una ingente cantidad de agua y electrolitos se vierten a la luz intestinal. Ello provoca diarreas y deshidratación. *V. cholerae* elabora una toxina que libera a la luz del intestino, reacciona con la mucosa intestinal y desencadena la secreción. La glucosa y los aminoácidos no sufren modificación en su absorción y además estimulan la absorción de agua y sales en la misma magnitud que

CORTE DE INTESTINO DELGADO, en representación esquemática, a la izquierda. La pared del intestino delgado consta de capas sucesivas de tejido muscular, tejido conectivo y epitelio simple, que es el que controla el tránsito de materiales (*ilustración superior*). Los enterocitos son células altas, columnares, cuya superficie apical (la que cae hacia la luz del intestino) presenta un gran número de microvellosidades, a manera de proyecciones digitales, de ahí el nombre de borde en cepillo (*esquemas inferiores*).



MECANISMO DE ABSORCION del cloruro sódico y agua. Se realiza a través de un proceso activo (bomba), señalado con una *S* invertida en la figura. La absorción de ClNa es consecuencia de la presencia en el borde en cepillo de un mecanismo que transloca Cl^- y Na^+. El Na^+ es bombeado hacia fuera de la célula, a través de la membrana basolateral, por la ATPasa Na^+ y K^+. Debido a este transporte activo de Na^+, el Cl^- se encuentra en el interior de la célula por encima de su potencial de equilibrio electroquímico y difunde hacia afuera por la membrana basolateral. El movimiento de agua sigue al de iones.



SECRECION INTESTINAL DE ClNa y H_2O, de acuerdo con el esquema propuesto por Field en 1979. El signo (+) indica estimulación y el menos (−), inhibición. En el dibujo de la pared intestinal (*derecha*), las evaginaciones se denominan vellosidades y las invaginaciones criptas de Lieberkühn. En la absorción de ClNa a través de una célula de epitelio que rodea las puntas de las vellosidades, los secretagogos inhibirían la entrada acoplada de ClNa por el borde en cepillo (*arriba, a la izquierda*). En el movimiento de ClNa a través de una célula epitelial situada en las criptas (*abajo, a la izquierda*), el Cl^- entra por la membrana basolateral de estas células acoplado al Na^+. El Na^+ es reciclado por la actividad del enzima ATPasa Na^+ y K^+. Los secretagogos estimularían la conductancia al Cl^- a través de la membrana mucosal.

en los individuos normales. Observación ésta que pone de manifiesto la selectividad de la alteración que provoca la toxina. Otros microorganismos, así *Clostridium welchii, Shigella dysenteriae* y *Escherichia coli*, producen toxinas capaces de provocar diarreas.

Hay muchas sustancias más que, al igual que la toxina del cólera, modifican el balance normal del transporte intestinal de agua y electrolitos dando lugar a una secreción neta. Recordemos la teofilina, el polipéptido vasointestinal (PIV), A23187, las prostaglandinas, la serotonina y el adenosín monofosfato cíclico (AMPc). Todas estas sustancias reciben el apelativo general de secretagogos. El que algunas de ellas las produzca el organismo (serotonina, polipéptido vasointestinal y prostaglandinas) y puedan ejercer su acción en condiciones fisiológicas ha despertado cierta expectación en los últimos años. Una misión verosímil de la secreción fisiológica de estas sustancias quizá fuera la de evitar una rápida deshidratación del contenido intestinal, que, en teoría, acabaría por interferir con una adecuada mezcla y digestión de los alimentos. Otro papel pudiera ser el de disminuir la velocidad de entrada de agua y sales, evitando así amplias oscilaciones en el volumen y la composición electrolítica del medio interno. Por lo que se refiere a otros agonistas naturales, las endomorfinas por ejemplo, actuarían aumentando la absorción.

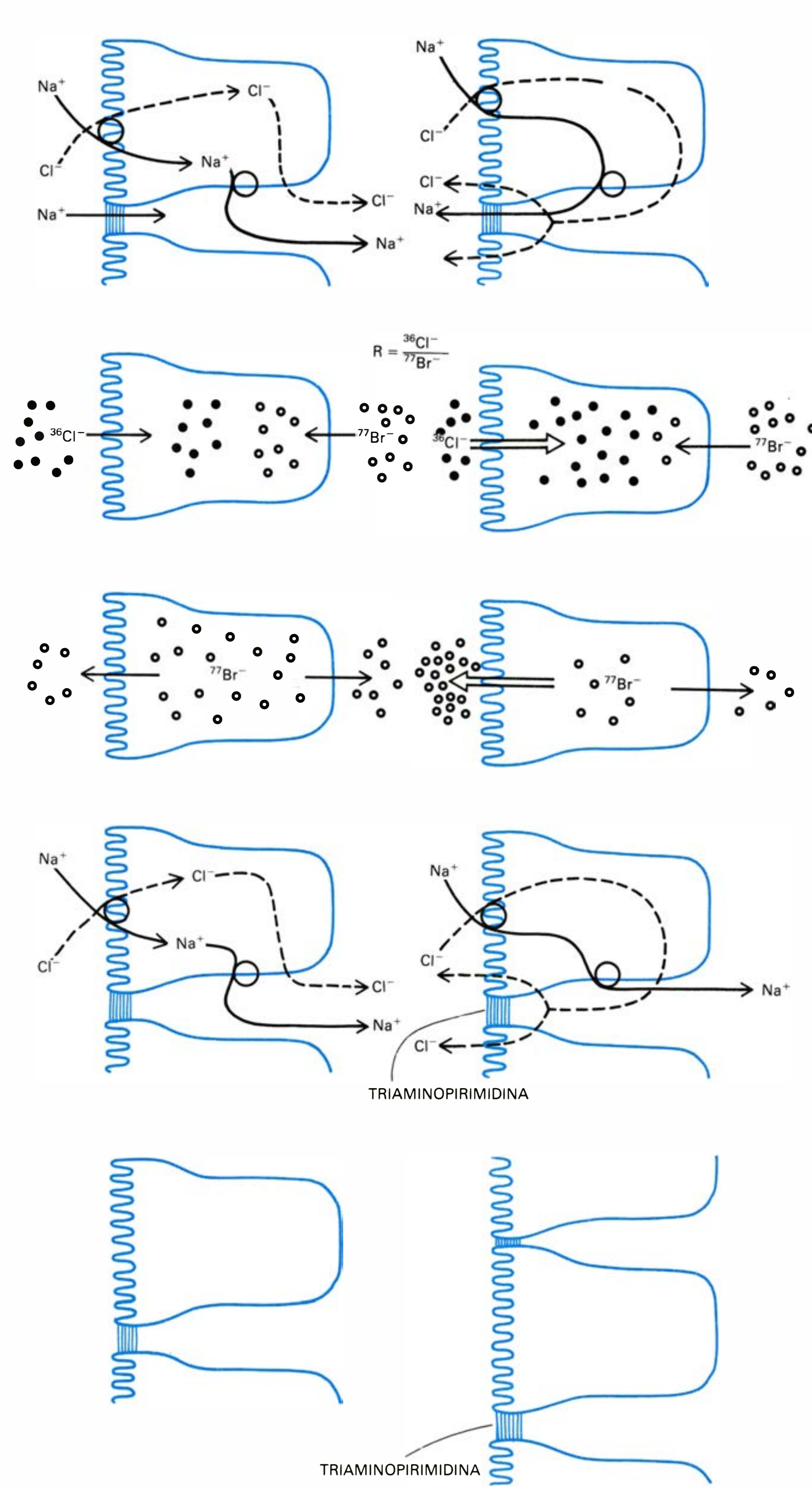
Desconocemos todavía los pormenores del mecanismo normal de control que mantiene ese delicado equilibrio entre procesos de absorción y secreción, ni se sabe qué parte asignar a las distintas poblaciones celulares encargadas de un proceso u otro. Lo que observamos al abordar el efecto de la teofilina, la toxina del cólera y el AMPc, sobre el movimiento de iones y parámetros eléctricos del intestino delgado del conejo, es una secreción neta de iones sodio y cloruro (experimentalmente reflejada en una disminución en el flujo de Cl^- y Na^+ desde el medio mucosal al serosal, sin cambio apreciable en los flujos en sentido contrario), acompañada de un aumento en la diferencia de potencial transmural.

A partir de esos resultados se han propuesto varias hipótesis para explicar la secreción intestinal desencadenada en respuesta a los secretagogos. En el primer tercio del decenio pasado, Powell, Binder y Curran adelantaron que la secreción de agua y electrolitos en íleon de conejo expuesto a secreta-

gogos debíase, en último término, a la estimulación en las células del epitelio de un proceso secretor, eléctricamente neutro, que transportara cloruro sódico e hidrocarbonato de sodio, o uno solo de ambos, desde el lado serosal hasta el mucosal. Contemporánea o inmediatamente después, Nellans, Frizzell y Schultz argumentaban que los secretagogos inhibían la entrada neutra acoplada de ClNa a través del borde mucosal (o borde en cepillo) de las células epiteliales; lo que dejaba al descubierto el mecanismo que media la salida del ClNa, que resulta enmascarado en condiciones normales por el proceso de entrada.

Field y sus colaboradores han venido defendiendo la idea de que los secretagogos intervendrían en dos procesos distintos implicados en el transporte intestinal de iones: en las células epiteliales situadas en las puntas de las vellosidades inhibirían la entrada neutra acoplada de ClNa y, en las células epiteliales de la base de las vellosidades, estimularían un proceso activo de secreción iónica. Este segundo proceso de secreción del Cl^- también estaría acoplada al Na^+ de la siguiente manera: por la membrana basolateral el Cl^- entraría en la célula acoplada al Na^+, pero este catión lo reciclaría la ATPasa Na^+ y K^+, en tanto que el Cl^-, en virtud del gradiente de sodio, se acumularía en la célula por encima de su equilibrio electroquímico y saldría por la membrana mucosal pasivamente, es decir, a favor de gradiente. Los secretagogos estimularían este proceso aumentando la permeabilidad al Cl^- de la membrana mucosal de estas células. El movimiento a través de los epitelios del Cl^- se acompañaría de un movimiento pasivo de Na^+ y agua.

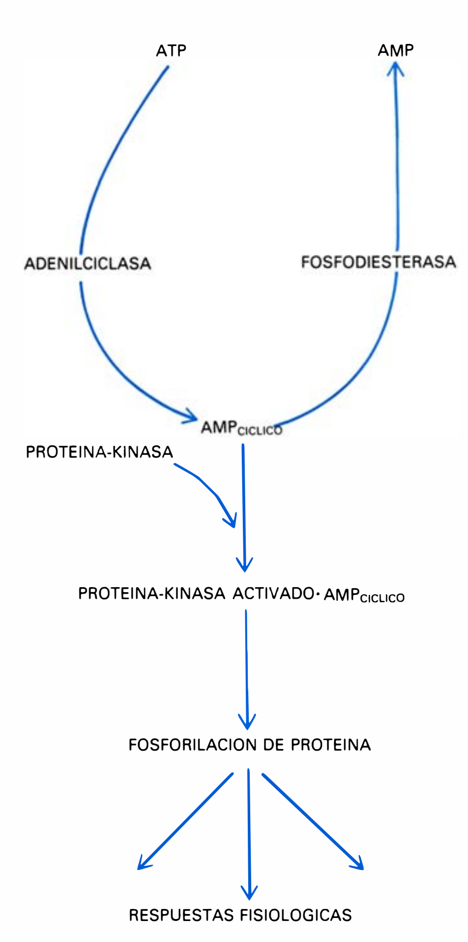
En 1979, otros autores (Naftalin y colaboradores) sugerían que los secretagogos ejercían una acción única en el íleon de conejo, a saber: la de aumentar la permeabilidad del Cl^- a través del borde en cepillo de las células del epitelio. Esquematizaban así el proceso: la dirección del movimiento neto de electrolitos a través del epitelio intestinal viene determinada por la relación de conductancia iónica entre uniones intercelulares —borde en cepillo, membrana basal— y lámina propia. Si la conductancia de la membrana mucosal superaba la de la membrana basal, se produciría una secreción neta de agua y electrolitos; en caso contrario, se daría absorción. Bajo condiciones normales, razonaba Naftalin, y debido



MODELO PROPUESTO POR NAFTALIN y colaboradores, en una secuencia de cinco pasos. En condiciones control, la permeabilidad de la membrana mucosal para aniones es menor que la de la membrana serosal, y el Cl^- y el Na^+, concentrados en los espacios intercelulares, difunden hacia el lado serosal. Cuando los secretagogos abren canales para el Cl^- a través del borde en cepillo, el Cl^-, acompañado de Na^+ y agua, sale hacia el lado mucosal desde los espacios intercelulares (*arriba*). El dibujo siguiente recoge el valor de *R*, es decir, de la relación en el tejido de actividades específicas de radioisótopos del cloro y bromo. Los secretatogos aumentan el valor de *R*. El esquema central ilustra la salida de Br^--77 desde el tejido. La triaminopirimidina, al bloquear el paso de Na^+ a través de las uniones intercelulares, disminuye en condiciones control la absorción de Na^+ y H_2O desde los espacios intercelulares, sin afectar el movimiento del Cl^- (*cuarto paso*). Los secretagogos cierran los espacios intercelulares (*abajo*).

a la actividad de la ATPasa Na^{+} y K^{+}, los iones sodio y cloruro quedan retenidos en los espacios intercelulares de íleon de conejo a concentraciones hipertónicas con respecto a la solución que baña el medio mucosal y serosal. A pesar de la alta conductancia catiónica de las uniones intercelulares, proseguía Naftalin, la permeabilidad aniónica a través de ellas es mucho menor que en el polo basal de los espacios intercelulares; por tanto, el volumen de ClNa que sale desde tales espacios hasta el lumen (lado mucosal) resulta despreciable.

Cuando los secretagogos abren canales para el Cl^{-} o aniones a través del borde en cepillo, aumentando así la conductancia aniónica de la membrana mucosal, se producirá una entrada pasiva de cloruro al medio mucosal desde los espacios intercelulares, donde el potencial electroquímico de dicho anión es más alto que en la solución mucosal o serosal. Ese reflujo de Cl^{-} irá acompañado por una salida de Na^{+} desde esos mismos espacios, a través de las uniones intercelulares, donde la conductancia del Na^{+} es siempre alta. La secreción pasiva o a favor de gradiente del ClNa comporta la salida de agua.



METABOLISMO y modo general de acción del adenosín monofosfato cíclico (AMPc). La adenilciclasa cataliza la formación de AMPc a partir de ATP y la fosfodiesterasa cataliza el paso a AMP. Un aumento de AMPc, bien porque se estimule la adenil-ciclasa o porque se inhiba la fosfodiesterasa, conduce a la activación de una proteína kinasa. La proteína kinasa activada fosforila determinadas proteínas, que actúan como efectores, que desencadenan la correspondiente respuesta celular.

El reflujo de Cl^{-} y Na^{+} (en relación 1:1) explicaría por qué los secretagogos disminuyen el flujo de Cl^{-} y Na^{+} desde el lado mucosal al serosal. Efectivamente, esta disminución podría ser debida: bien a que se inhiba el mecanismo responsable de la translocación del par iónico, Cl^{-} y Na^{+}, a través del borde en cepillo, tal y como postula Field, o bien a que este mecanismo no se altere, pero se estimule la salida de Cl^{-} y Na^{+} (reflujo), con lo que el resultado neto será una disminución en la entrada.

Los datos experimentales en que se funda el modelo se recabaron al estudiar, en íleon de conejo, el efecto de los secretagogos teofilina y toxina del cólera sobre distintos parámetros: flujo pasivo de Cl^{-} y Na^{+} a través del epitelio, actividades específicas en el tejido de los isótopos radiactivos del cloruro (Cl-36) y bromuro (Br-77), salida pasiva de bromuro desde el tejido y, por último, empleo de triaminopirimidina (TAP). Vamos a detallarlos.

El flujo pasivo de Cl^{-} y Na^{+} por el epitelio, en presencia de ouabaina y con adecuados gradientes iónicos, aumenta con la acción de los secretagogos. En presencia de ouabaina, cualquier modificación en el flujo de un ion se deberá sólo a cambios ocurridos en la permeabilidad del tejido.

La teofilina y la toxina del cólera incrementan el valor de la relación *R*: relación en el tejido de actividades específicas de radioisótopos del cloruro (Cl-36) y bromuro (Br-77), procedentes del lado mucosal y serosal, respectivamente. Si los secretagogos actuaran inhibiendo la entrada neutra de ClNa y estimulando el proceso secretor, como decían otros autores, el valor de *R* en el tejido secretor sería menor que en el control. En efecto, entraría en el primero menos cloruro desde el lado mucosal y más bromuro desde el serosal. Al no incidir en ese aumento la ouabaina, sugiere de nuevo que los cambios provocados en el movimiento de Cl^{-} por los secretagogos son cambios en las propiedades de permeabilidad del ion a través del borde en cepillo.

La teofilina intervenía en la salida pasiva, en presencia de ouabaina, del ion bromuro desde el tejido cargado previamente con Br-77. El secretagogo aumentó la tasa de salida del Br-77 a la solución mucosal y redujo la del ion al lado serosal.

El empleo de triaminopirimidina (TAP) permite revelar la existencia de diferentes canales para el paso de Na^{+} y Cl^{-} a través de la membrana celular. Se sabe que la TAP bloquea, en un gran número de epitelios, el movimiento de cationes monovalentes a través de las uniones intercelulares, pero no afecta el movimiento del Na^{+} a través del borde en cepillo. En nuestro sistema secretor, TAP bloqueó el efecto de la teofilina sobre el movimiento del sodio, sin modificar el valor de *R* obtenido para el Na^{+} o el Cl^{-} y sin afectar el flujo de Cl^{-}. Estos resultados revelan que la secreción de Na^{+} no se debe exclusivamente a una inhibición de su entrada en las células, ya que si así fuera la TAP no tendría efecto alguno sobre la secreción del catión, e indican, asimismo, que durante el estado secretor la salida de Na^{+} ocurre a través de las uniones intercelulares. Cuando quedan éstas bloqueadas, no puede salir Na^{+} hacia el lado mucosal.

Todos los resultados favorecen la hipótesis de que la secreción intestinal de agua y electrolitos, en respuesta a secretagogos, obedece, de forma exclusiva, a un aumento en la conductancia de Cl^{-} a través del borde en cepillo. El mecanismo responsable de la translocación acoplada de ClNa, localizado en el borde en cepillo, no se ve afectado por los secretagogos. (Por translocación acoplada se entiende el paso de dos iones de carga opuesta en tal relación que el proceso sea eléctricamente neutro. Así en el caso del Cl^{-} y Na^{+} el mecanismo transloca 1 Cl^{-} por cada 1 Na^{+}).

El microscopio electrónico pone de manifiesto que la teofilina cierra los espacios intercelulares, acción que queda en suspenso en presencia de TAP. Lo que concuerda con la opinión que afirma que la salida de Na^{+} y Cl^{-}, a favor de gradiente desde los espacios intercelulares, produce un arrastre electroosmótico de agua desde esos espacios. Ese fenómeno anatómico observado en presencia de la teofilina no se da sólo en la parte del epitelio localizado en las puntas de las vellosidades, sino que se extiende hasta la base de las mismas y región de las criptas, lo que presta soporte a la hipótesis de que es una misma población celular la responsable del fenómeno de absorción y del fenómeno de secreción.

¿Qué modificación se produce en el interior de las células, tras interaccionar con ellas los secretagogos, que in-

duce un cambio en la dirección neta del movimiento de iones y agua? Dicho de otra manera: ¿Cuál es la naturaleza del factor de acoplamiento o segundo mensajero entre la estimulación y la respuesta celular consiguiente?

Son varios los segundos mensajeros a los que se ha implicado en el control del movimiento de iones. Además del calcio, que parece desempeñar un papel central en muchos sistemas celulares, se citan también dos nucleótidos cíclicos: AMPc y GMPc. El nivel intracelular de los nucleótidos cíclicos está determinado por el balance dinámico entre el proceso de síntesis, catalizado por la adenil ciclasa o guanidil ciclasa, respectivamente, y el de degradación, catalizado por las correspondientes fosfodiesterasas. Cualquier agente que estimule la adenil ciclasa o la guanidil ciclasa o inhiba la fosfodiesterasa producirá un aumento del nivel intracelular del correspondiente nucleótido cíclico.

En las células de mamíferos, la mayoría de las acciones de los nucleótidos cíclicos están mediadas por proteínas kinasas, que, al unirse con el correspondiente nucleótido, se activan y producen la fosforilación de proteínas. Estas proteínas fosforiladas sirven de efectores o moduladores de reacciones celulares o respuestas fisiológicas.

A diferencia del metabolismo celular de los nucleótidos cíclicos, el metabolismo celular del calcio es algo complejo. Este catión se encuentra desigualmente distribuido por los compartimentos celulares. En cualquier momento, la concentración citosólica del calcio ionizado, $[Ca^{2+}]$, resulta del equilibrio dinámico entre mecanismos que regulan la entrada y salida del mismo a través de la membrana plasmática de las células y de la membrana de compartimentos intracelulares que secuestran calcio (mitocondrias y retículo sarcoplasmático, sobre todo) y mecanismos citosólicos que amortiguan Ca^{2+} (componentes citoplasmáticos que acomplejan calcio). Se supone que la concentración de calcio ionizado en el citoplasma ronda los 10^{-7} milimol $- 5 \times 10^{-8}$ milimol, mientras que en la matriz mitocondrial es de $10^{-5} - 10^{-4}$ milimol y en el medio extracelular (en mamíferos) es del orden de 10^{-3} milimol. El contenido citosólico de Ca^{2+} es el principal responsable de las funciones del Ca^{2+} como segundo mensajero.

Un aumento en el nivel citoplasmático de calcio ionizado hasta valores de $10^{-6} - 5 \times 10^{-6}$ milimol facilita cambios funcionales que, según el tipo celular de que se trate, determina múltiples

MEDIO EXTRACELULAR
CITOPLASMA
MEMBRANA PLASMATICA
CALCIO ACOMPLEJADO POR ESTRUCTURAS CITOSOLICAS
[CALCIO] 10^{-3}MOLAR
[CALCIO] 10^{-7}MOLAR
CALCIO EN MITOCONDRIA
[CALCIO] 10^{-5} MOLAR
(+)
ESTIMULACION CELULAR
CALCIO EN RETICULO ENDOPLASMATICO
[CALCIO] 10^{-6}MOLAR
Ca++
PRC
PROTEINAS RECEPTORAS DEL CALCIO
Ca++
PRC
FORMA ACTIVADA
ENZIMA
ENZIMA ACTIVADO
E
Ca++
PRC
REACCIONES BIOQUIMICAS
RESPUESTAS FISIOLOGICAS

METABOLISMO Y MODO DE ACCION del calcio. A diferencia del metabolismo celular de los nucleótidos, el del calcio es bastante complejo. Este ion se encuentra desigualmente distribuido por los compartimentos celulares. En cualquier momento, la concentración citosólica del calcio ionizado resulta del balance dinámico entre mecanismos que regulan la entrada y salida del calcio a través de la membrana plasmática de la célula y de las membranas de los compartimentos intracelulares. Los círculos representan el mecanismo que activamente trasloca Ca^{2+} a través de las membranas. En ausencia de este signo, las flechas indican movimientos pasivos. La estimulación de la célula conduce a un aumento de la concentración de calcio ionizado en el citoplasma hasta valores de 10^{-6} molar o más, lo que permite la formación del complejo Ca-PRC. Este complejo es la forma activa y regula la actividad de un enzima apropiado.

respuestas; entre ellas: contracción muscular, exocitosis en glándulas endocrinas y exocrinas, iniciación de la síntesis de ADN, movimiento cromosómico durante la mitosis, motilidad celular, alteración de la permeabilidad de la membrana plasmática a iones y otras.

Todas estas observaciones apoyan la hipótesis que el calcio funciona como segundo mensajero en los sistemas biológicos, en forma análoga a como lo hace el AMPc. Mientras que el principal mediador del efecto regulador del AMPc es una kinasa, el factor celular que media la acción reguladora del calcio es menos conocido. Durante los últimos años los investigadores han centrado sus esfuerzos en aislar proteínas que forman complejos específicos con calcio con objeto de identificar el receptor o los receptores intracelulares del mencionado catión. La información disponible al respecto sugiere que los diferentes efectos del calcio vienen mediados por proteínas que actúan de receptores intracelulares del calcio, en analogía a la función receptora que cumple la kinasa con los correspondientes nucleótidos cíclicos. Numerosas pruebas sugieren que la estimulación de las células conduce a un aumento transitorio de la concentración citoplasmática de calcio ionizado hasta niveles suficientes para provocar la unión del Ca^{2+} con las proteínas receptoras. Ni el calcio por sí solo, ni la proteína receptora logran activar los sistemas enzimáticos. Pero la proteína receptora sufre cambios en su configuración tras incorporar el calcio en su molécula. El complejo formado por la proteína y el calcio se une a una proteína efectora que desencadena una reacción bioquímica, que culmina en una respuesta celular.

En relación con la secreción en el intestino delgado, en respuesta a la presencia de secretagogos, se ha postulado la existencia de dos segundos mensajeros celulares. La acción secretora de determinados agonistas, tales como la teofilina, la toxina del cólera y el VIP, se acompaña de un aumento de la concentración de AMP cíclico intracelular, pero no depende de la presencia de Ca^{2+} en el medio extracelular. Esto indujo a pensar que en la acción de esos secretagogos intervenía el AMPc. Un aumento en la concentración intracelular de AMPc conduciría a la activación de una determinada proteína kinasa, que fosforilaría determinada proteína de membrana y provocaría así cambios en la permeabilidad iónica a través de la membrana plasmática.

Otros secretagogos (serotonina, carbamilcolina y A23187) no modifican el nivel intracelular de AMPc. Su acción depende de la presencia de calcio extracelular, lo que permite adjudicar al calcio el papel de segundo mensajero en la acción de estos agonistas.

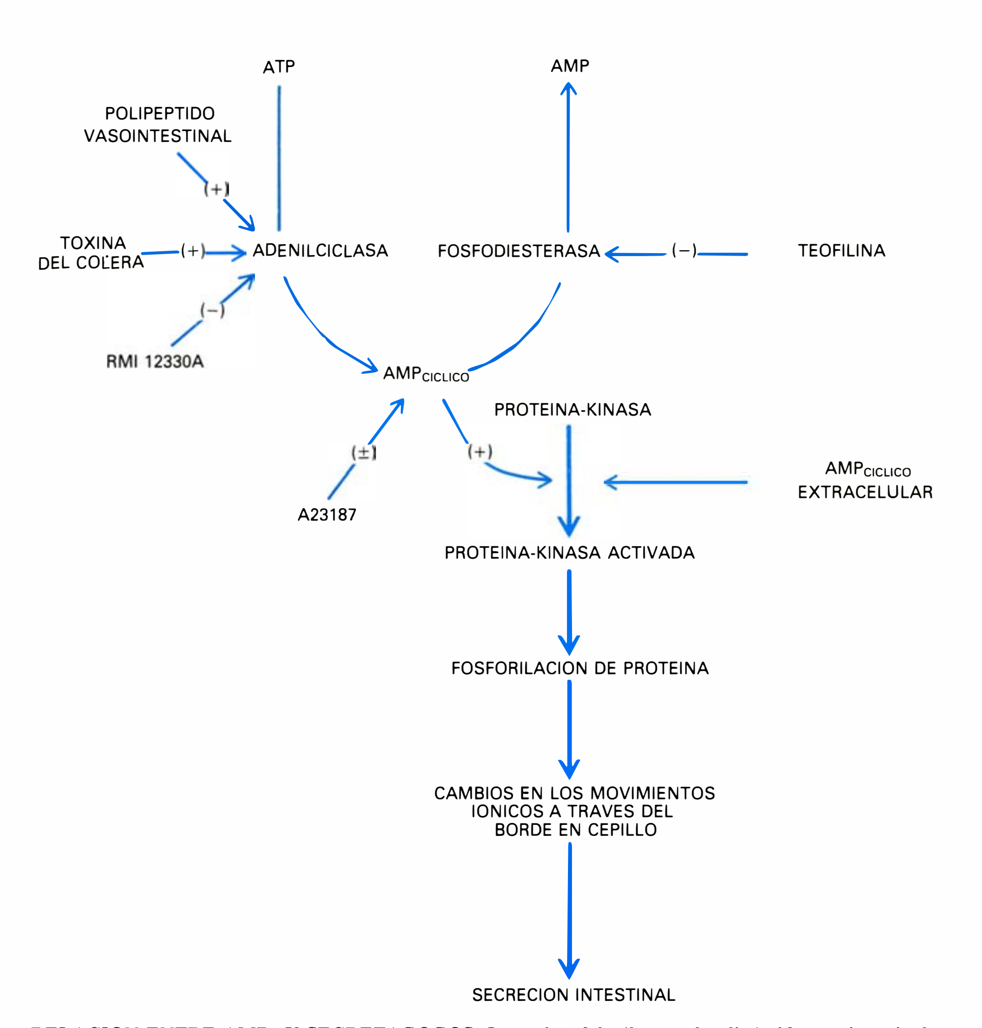
EFECTO DE SECRETAGOGOS, RMI 12330A Y ESTELAZINA sobre los flujos de cloro a través de íleon de conejo. Los flujos de cloro se miden en las cámaras. Cada cámara consta de dos semicámaras, dotadas, a su vez, de seis pocillos y cada pocillo presenta dos orificios (uno apical y otro lateral). Cuando se juntan las dos semicámaras los 12 pocillos se comunican dos a dos lateralmente. Se abre longitudinalmente el íleon de conejo y se coloca entre las dos semicámaras. En cada pocillo se pone solución de Ringer, sin (condiciones control) o con modificadores, se gasea con 95 por ciento de oxígeno y cinco por ciento de dióxido de carbono y se mantiene a 37º C. Para medir el flujo de Cl^- desde mucosal a serosal el Cl^--36 se añade al lado mucosal y para el flujo inverso el Cl^--36 se añade al serosal. Se toman muestras de los pocillos que inicialmente carecían de radioisótopo. Los resultados aparecen en el dibujo inferior.

Los resultados obtenidos por nosotros al estudiar el efecto de trifluoroperazina (tranquilizante del grupo de las fenotiazinas, también llamado estelazina) y de RMI 12330A (una lactanamida), sobre la acción de los secretagogos en íleon de conejo, nos permitió

llegar a la conclusión de que, en los propios casos en los que se elevaba el nivel intracelular de AMPc, la secreción no podía ser desencadenada por la acción del AMPc sobre la membrana plasmática. Se ha demostrado que RMI 12330A es un inhibidor de la adenil ciclasa en células hepáticas y se ha demostrado que inhibe, en colon humano, la secreción inducida por VIP y prostaglandinas. La estelazina, al igual que otros tranquilizantes, inhibe específicamente una de las múltiples formas de fosfodiesterasa aislada de cerebro bovino, cuya actividad depende asimismo del Ca^{2+}.

En nuestro sistema secretor, la estelazina y la RMI 12330A bloquearon el efecto que los secretagogos tienen sobre el movimiento de Cl^- Ninguno de estos fármacos afectó el movimiento de Cl^- en ausencia de secretagogos, lo que puso de relieve que ambos actúan sólo cuando los secretagogos han estimulado las células y en consecuencia se ha modificado algún componente intracelular en respuesta a ellos.

De acuerdo con su calidad de inhibidor de la adenil ciclasa, RMI 12330A se interpuso al aumento de AMPc provocado por la teofilina (inhibidor de la fosfodiesterasa) o por la toxina del cólera o el VIP (estimuladores de la adenil ciclasa). RMI 12330A, sin embargo, bloqueó también la secreción de Cl^- desencadenada por AMP cíclico extracelular (AMPc añadido al medio de incubación). Si RMI 12330A actuara únicamente como inhibidor de la adenil ciclasa y la acción del AMPc, mediada por la correspondiente kinasa, se ejerciera sobre la membrana de las células, RMI 12330A no debería bloquear la acción secretora del AMPc extracelular, ya que la RMI intervendría en la etapa anterior a la formación de AMPc celular, que sería redundante en presencia del AMPc extracelular. Además, RMI 12330A bloqueó igualmente la secreción de Cl^- en aquellas condiciones en las que el nivel intracelular de AMPc no se modificaba, por ejemplo en respuesta al ionóforo A23187 (transportador del Ca^{2+}). Estos resultados revelaron que RMI 12330A, además de inhibir adenil ciclasa, bloqueaba alguna otra reacción bioquímica intracelular que se desencadenaba tanto en respuesta a los secretagogos que aumentaban el nivel celular de AMPc como en respuesta a los agonistas que no modificaban la concentración celular del nucleótido cíclico. Por otro lado, la estelazina no modificó los niveles de AMPc observados en el tejido expuesto a secretagogos o en los tejidos control.



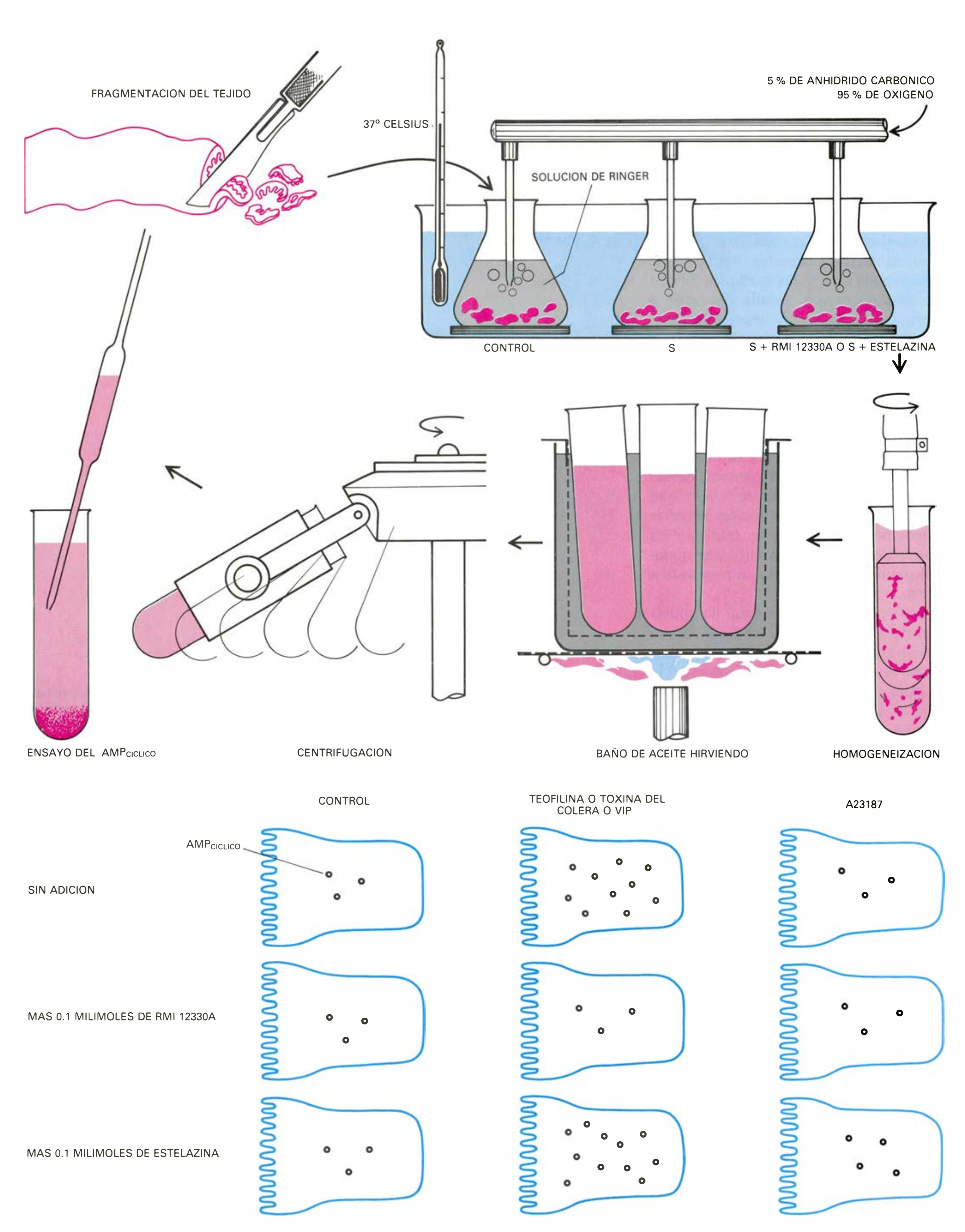
RELACION ENTRE AMPc Y SECRETAGOGOS. La toxina del cólera y el polipéptido vasointestinal (VIP), al estimular la adenil-ciclasa (AC), aumentan la concentración intracelular de AMPc. La teofilina produce el mismo efecto inhibiendo la fosfodiesterasa (FD). La RMI 12330A, al inhibir la AC, bloquea el efecto de la teofilina, la toxina del cólera o el VIP sobre el nivel de AMPc. El A23187 no modifica los niveles de AMPc. Para los autores, un aumento de la [AMPc] celular activaría una proteína binaria, la cual fosforilaría determinadas proteínas de membrana, provocando así cambios en los movimientos iónicos a través de la membrana de los enterocitos, lo que daría lugar a secreción de Cl^-, Na^+ y agua.

Quedaba así de manifiesto que el aumento en el nivel citoplasmático de AMPc, por sí solo, no era capaz de alterar la permeabilidad aniónica de la membrana citoplasmática, ya que la secreción intestinal ocurría tanto cuando la concentración del nucleótido en el tejido era alta (caso de la teofilina o de la toxina del cólera) como cuando era baja (en respuesta al A23187). Además, la estelazina bloqueaba siempre la secreción sin afectar los niveles intracelulares del nucleótido.

Llegamos a resultados similares tras investigar el efecto de secretagogos, estelazina y RMI 12330A, sobre el nivel intracelular de GMPc. Así descartamos al GMPc como mediador intracelular de la secreción intestinal.

Si la variación en la concentración intracelular de nucleótidos cíclicos no es, por sí sola, el factor de acoplamiento entre estimulación y secreción intestinal, ¿cuál es la naturaleza del mensajero intracelular que liga estos mecanismos?

Ya hemos mencionado que otro componente de los sistemas de señalización celular es el calcio. El modo más directo de demostrar que el calcio constituye el factor puente entre estímulo-secreción es el de probar que, durante el acoplamiento estímulo-secreción, se produce un aumento en la concentración citoplasmática de calcio ionizado, de forma análoga a como ocurre en la célula muscular durante el acoplamiento excitación-contracción. A diferencia de las células musculares, las secretoras son demasiado pequeñas para valernos del uso de acuarina, sustancia que, inyectada en la célula, detecta cambios en la actividad citoplasmática de calcio. En estos casos el uso



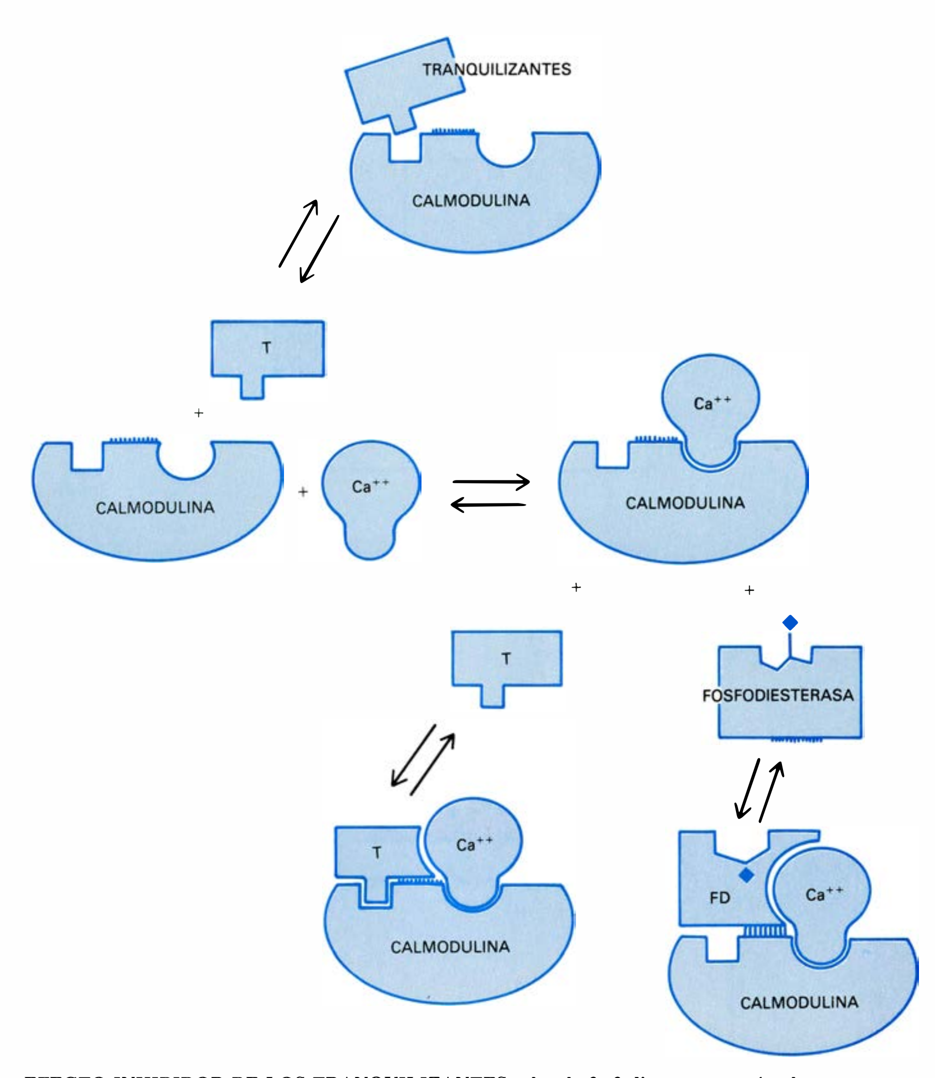
EFECTO DE LOS SECRETAGOGOS, la RMI 1233OA y la estelazina sobre el nivel de AMPc en el tejido. En el experimento ideado por los autores, el tejido se cortaba en trocitos. Los trocitos se repartían en diferentes recipientes que contenían solución de Ringer (con o sin drogas), se colocaban en un baño de agua mantenida a 37 grados C y se incubaban cierto tiempo con continuo burbujeo de carbógeno. Terminado el período de incubación cada tejido fue homogenizado en solución de EDTA, hervido en baño de aceite y luego centrifugado. Se tomaron muestras del sobrenadante. Resultó lo siguiente: la RMI 12330A bloqueó el aumento de AMPc en respuesta a la teofilina, toxina del cólera o VIP y la estelazina no modificó el nivel de AMPc en el tejido.

de microelectrodos permite controlar cambios en el nivel intracelular de calcio ionizado. Otro método de demostrar la importancia del calcio es el empleo de ionóforos del calcio. De entre todos los ionóforos de cationes divalentes conocidos hasta el momento, el llamado A23187 se ha convertido en el más específico y útil para el Ca^{2+}. Se desconoce el mecanismo preciso por el que el A23187 aumenta la permeabilidad del Ca^{2+} a través de las membranas de las células. Se cree que interacciona con las membranas y que, por cada dos moléculas de A23187, se forma un poro a través del cual el Ca^{2+} puede atravesarlas.

Se ha visto que el A23187 simula el efecto de agonistas naturales (5-hidroxitrictamina, acetilcolina, etcétera) en un gran número de sistemas secretores: glándulas salivares de insectos y mamíferos y córnea de rana, entre otros. Como ya se ha indicado anteriormente, simula la acción de la toxina del cólera en íleon del conejo.

La toxina del cólera y el ionóforo A23187 desencadenan en íleon de conejo la misma respuesta celular. ¿Significa ello que la toxina del cólera y los otros secretagogos, cuya acción se acompaña de un aumento en el nivel intracelular de AMPc y es independiente de la presencia de calcio en el medio de incubación, aumentan la concentración citosólica de calcio ionizado?

Con el fin de detectar si la actividad citosólica de Ca^{2+} se modificaba durante el estado secretor, con independencia del agonista empleado para desencadenar la secreción, recurrimos a un método indirecto. Consistía en medir la entrada en el tejido de estelazina marcada con tritio (un inhibidor de la secreción intestinal).

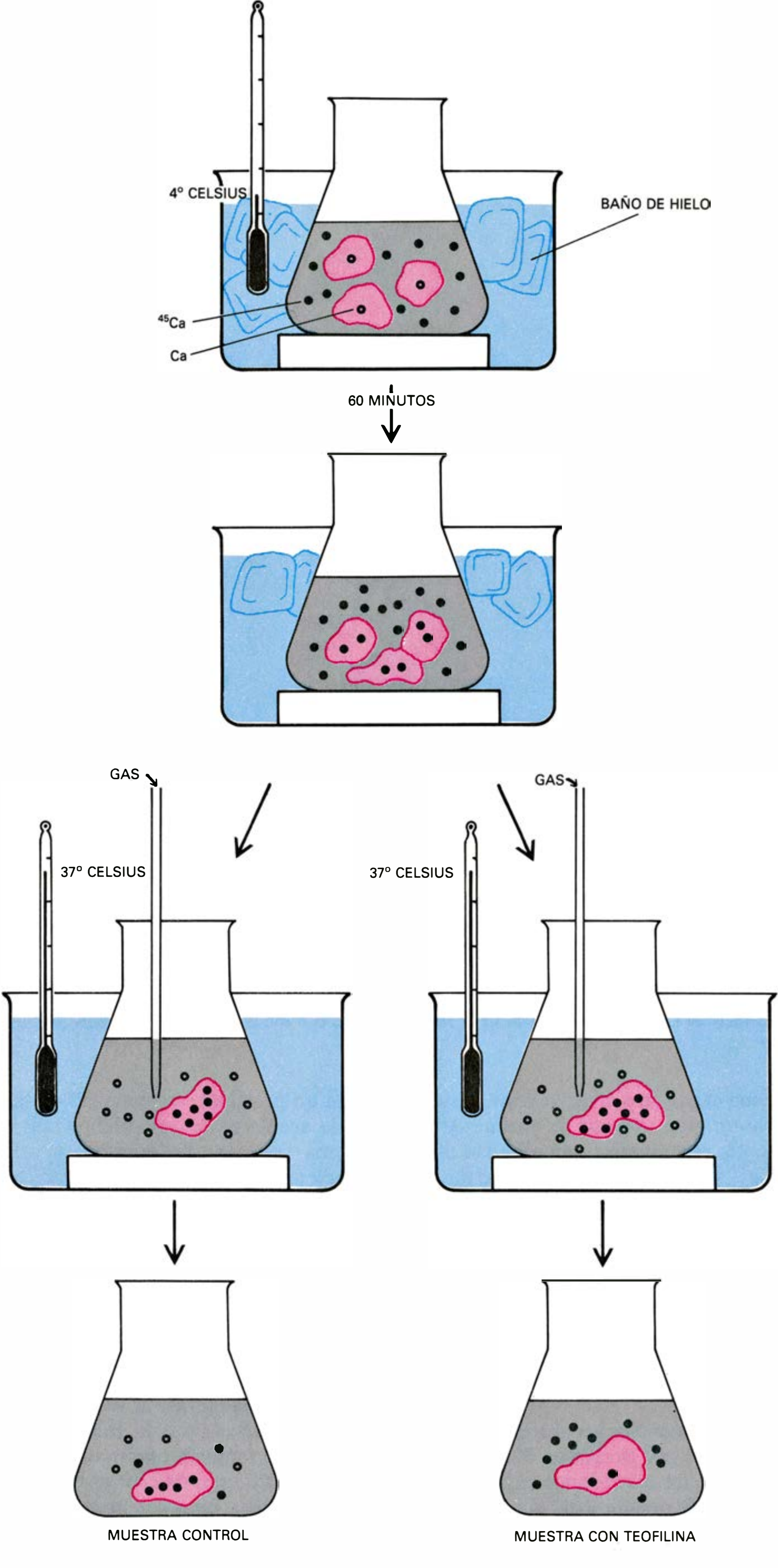


EFECTO INHIBIDOR DE LOS TRANQUILIZANTES sobre la fosfodiesterasa, según el esquema propuesto por Levin y Weiss. La primera etapa es la interacción de Ca^{2+} con la calmodulina (CDR), lo que produce un cambio conformacional de la calmodulina. El complejo Ca-CDR interacciona con una forma molecular específica de fosfodiesterasa (FD) aislada de cerebro bovino y se forma el complejo activo Ca-CDR-FD. Alternativamente, el complejo Ca-CDR se puede unir a agentes tranquilizantes (T) formando el complejo Ca-CDR-T. Este complejo no puede estimular la forma activable de fosfodiesterasa. En ausencia de Ca^{2+} la interacción de CDR con fosfodiesterasa o con el tranquilizante es muy pequeña.

¿Por qué este experimento podría revelarnos si los secretagogos modificaban los niveles citosólicos de Ca^{2+}? Los estudios realizados en relación con el modo de acción del calcio muestran que de entre todas las proteínas celulares, identificadas hasta el momento, capaces de formar complejos con calcio (troponina C, pavalbúmina, S-100), la calmodulina, o CDR, parece ser el receptor intracelular más universal del calcio [véase "Calmadulina", por Wai Yiu Cheung; INVESTIGACION Y CIENCIA, agosto]. La calmodulina se distingue por varios aspectos: se encuentra en el citoplasma de muchos, si no todos, los eucariotas; ha conservado su estructura y función a lo largo de la escala vegetal y animal; es la única proteína capaz de activar de forma reversible un gran número de enzimas y presenta, por último, gran afinidad por el Ca^{2+} (constante de disociación de 10^{-6} M). Rasgos que sugieren que la calmodulina regula procesos constitutivos, comunes a muchos tipos de células. A diferencia de ella, las demás proteínas que incorporan el calcio con gran afinidad manifiestan una distribución más limitada y regulan funciones más especializadas y de extensión más restringida.

Son numerosos los procesos celulares regulados por la calmodulina: metabolismo de los nucleótidos cíclicos, fosforilación de membranas, unión y separación de microtúbulos, metabolismo del glucógeno y flujo de calcio entre otros. Para todo ello la calmodulina debe estar unida al calcio formando el complejo calmodulina-calcio (CDR-Ca).

En un principio, se relacionó el modo de acción de la calmodulina con el sistema de la fosfodiesterasa. El complejo Ca-CDR estimula una de las varias formas de fosfodiesterasa de cerebro bovino. La activación de ésta fosfodiesterasa por el complejo Ca-CDR es selectivamente inhibida por diversos tranquilizantes, entre ellos la estelazina. R. M. Levin y B. Weiss demostraron que la estelazina ejercía su efecto inhibidor uniéndose al complejo Ca-CDR, bloqueando así la activación de la fosfodiesterasa por la calmodulina. Observaron también que la calmodulina poseía dos tipos de centros de unión para estelazina, unos con gran afinidad (constante de disociación, K_D, del orden de 10^{-6} M) y con unión dependiente de Ca^{2+}, y los otros centros de unión con afinidad menor ($K_D = 5 \times 10^{-3}$ M) y con unión independiente de Ca^{2+}.



EFECTO DE LA TEOFILINA sobre la velocidad de salida de calcio desde el tejido previamente cargado con Ca-45. Cortábamos en trocitos el íleon de conejo. Para cargarlos con Ca-45 se colocaban en solución de Ringer que contenía Ca-45, mantenida a 4 grados. En presencia de la teofilina la velocidad inicial de salida de Ca-45 desde el tejido al medio de incubación fue mayor que la observada en los tejidos control.

A bajas concentraciones de estelazina, la relación entre la cantidad de estelazona unida al complejo Ca-CDR, con respecto a la unida a CDR, es del orden de 20:1. Conforme se aumenta la concentración de estelazina, disminuye la proporción, hasta hacerse nula. En presencia de Ca^{2+} la unión de estelazina a la calmodulina es específica: la estelazina sólo se une con gran afinidad al complejo CDR-Ca. Su unión a otras proteínas (albúmina de huevo, citocromo C, aldolasas, etcétera) es independiente de calcio y resulta similar a la observada con la CDR en ausencia de calcio.

Basándonos en estos datos nos hicimos el siguiente razonamiento: si los secretagogos aumentaban la concentración citoplasmática de calcio ionizado y si había calmodulina en las células del epitelio intestinal de íleon de conejo, un aumento en la concentración de Ca^{2+} en respuesta a los secretagogos conduciría a la formación del complejo Ca-CDR, al que la estelazina se unía con gran afinidad. En consecuencia, la cantidad de estelazina secuestrada en los tejidos expuestos a los secretagogos sería mayor que la secuestrada por los tejidos control, en los que lógicamente la cantidad de Ca-CDR será muy pequeña.

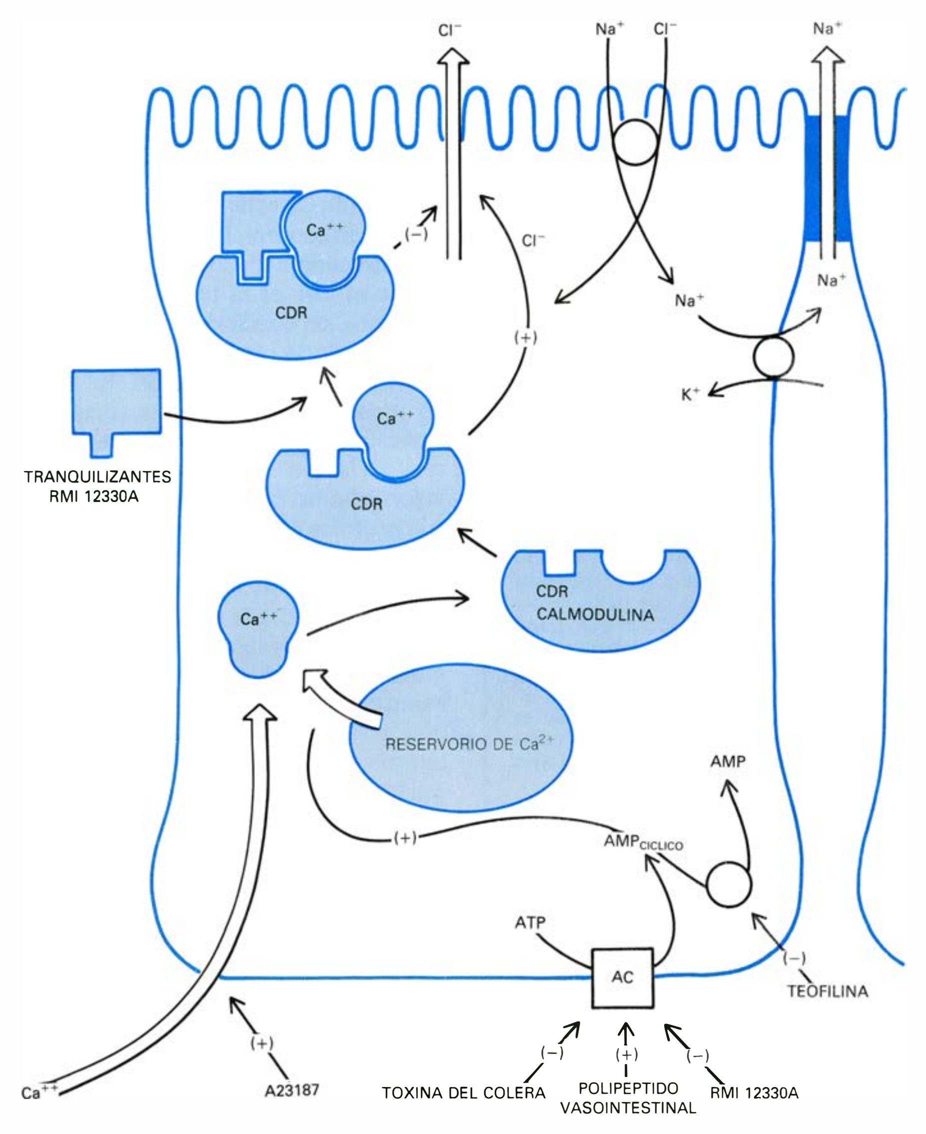
Los resultados obtenidos en estos experimentos corroboraron nuestra hipótesis. El A23187 estimuló la entrada de estelazina con respecto a los tejidos control. Incremento que se vio bloqueado cuando la actividad del calcio en el medio de incubación decayó por adición de EGTA, un quelante de calcio. La teofilina y la toxina del cólera estimularon la entrada de estelazina, lo que demostraba que, en presencia de estos secretagogos, además de producirse un aumento en el nivel citoplasmático de AMPc, también provocaban un aumento en la concentración de calcio citosólico. La RMI 12330A inhibió el aumento en la entrada de estelazina en respuesta a los secretagogos, lo que sugería que RMI 12330A competía por los centros intracelulares a los que se unía estelazina y que ambas, estelazina y RMI 12330A, bloqueaban la secreción inhibiendo una misma reacción bioquímica intracelular. Esto permitiría explicar por qué la RMI 12330A era capaz, como se recordará, de bloquear el efecto secretor del AMPc extracelular, o de aquellos secretagogos que no modificaban el nivel citoplasmático de AMPc (el A28187, por ejemplo). Otro tipo de experimentos, en los que se midió la entrada en el tejido de diferentes

concentraciones de estelazina, revelaron que la constante de disociación de estelazina por los centros intracelulares de íleon de conejo, estimulado por la teofilina el A23187, era bastante alta (K_D : 10^{-6} molar).

El hecho de que la acción de la toxina del cólera o la teofilina no dependiera de la presencia de Ca^{2+} extracelular y, sin embargo, estimularan, al igual que el A23187, la entrada de estelazina en el tejido, abogaba por la existencia de almacenes intracelulares de calcio que potencialmente aportarían todo el calcio necesario para mantener la secreción intestinal, sobre los que incidirían los secretagogos cuya acción se acompañaba de un aumento de AMPc celular. Ese hecho era, además, independiente de calcio extracelular y aumentaba la [Ca^{2+}] intracelular, aumento reflejado por una mayor entrada en el tejido de estelazina.

En algunos tipos de células secretoras se ha sugerido que el AMPc actúa liberando calcio de lugares intracelulares en donde normalmente se almacena. No sabemos gran cosa de la naturaleza de estos reservorios intracelulares de calcio sensibles al AMPc. Una forma indirecta de averiguar si durante la activación celular se libera calcio al citoplasma desde los reservorios (pools) intracelulares consiste en medir el efecto de agonistas sobre la salida de calcio desde las células hasta el medio extracelular. Para ello, primero se incuba el tejido en un medio que contiene Ca-45, lo que permite que dicho isótopo entre en los reservorios intracelulares que contienen calcio intercambiable. Terminado el período de incubación, se coloca el tejido en un medio carente de Ca-45 y se mide luego la salida de calcio desde el tejido a lo largo del tiempo, flujo que probablemente refleja la velocidad con que el calcio abandona los reservorios intracelulares. Si la estimulación del tejido por un determinado agonista conduce a un aumento en la velocidad de salida de Ca-45, se considera que se ha estimulado la velocidad a la que se libera calcio de los reservorios al citoplasma, en donde actúa como sustrato de la bomba de calcio, que lo saca al exterior. Este efecto se ha descrito en páncreas, glándulas salivares de insectos y mamíferos, hígado y en intestino grueso de mamífero. En íleon de conejo, utilizando este mismo método, observamos que la teofilina aumentaba la velocidad de salida de calcio desde el tejido hasta el medio de incubación cuando se comparaba con los tejidos control. Este resultado parece indicar que el aumento en la actividad citoplasmática de calcio observado en presencia de la teofilina, y medido por el aumento en la entrada de estelazina al tejido, posee un origen intracelular.

Hasta el momento todos nuestros resultados apoyan la hipótesis de que el Ca^{2+} es el mediador intracelular, común, en el proceso de secreción intestinal desencadenado en respuesta a los secretagogos. Sin embargo, como se recordará, la secreción intestinal de Cl^- desencadenada por el A23187 fue bloqueada por la estelazina y la RMI 12330A, lo que claramente demuestra que un aumento en la actividad citoplasmática de calcio, por sí solo, no estimula directamente la secreción. La acción del calcio debía estar mediada por algún receptor intracelular, que tras unirse al calcio, estimulaba la permeabilidad del Cl^- a través del borde en cepillo de las células del epitelio intestinal, lo que en último término desencadenaba la secreción intestinal. Los resultados obtenidos al estudiar la entrada de estelazina al tejido (recordemos la alta afinidad y especificidad que la estelazina, en presencia de Ca^{2+} tiene por CDR aislada de cerebro bovino) sugerían, por su similitud a los previamente obtenidos por otros autores,



MECANISMO PROPUESTO para explicar la secreción de Cl^- a través de íleon de conejo. El signo (+) representa activación y el (−), inhibición. Los secretagogos incrementan la concentración citosólica de Ca^{2+}, aumentando la permeabilidad al Ca^{2+} de la membrana plasmática (así el A23187) o a través de un aumento en la concentración citoplasmática de AMPc (teofilina y toxina del cólera por ejemplo). El AMPc actuaría sobre los almacenes intracelulares de Ca^{2+} provocando la liberación de éste al citoplasma. Un aumento en la concentración intracelular del calcio ionizado conduciría a la formación del complejo Ca-CDR, el cual actuaría sobre el borde en cepillo, abriendo canales para el Cl^-. La salida de Cl^- se acompaña de una salida de Na^+. Los tranquilizantes y la RMI 12330A inhibirían el efecto secretor de los secretagogos uniéndose al complejo Ca-CDR y bloqueando su acción sobre el borde en cepillo.

que el receptor del calcio en nuestro sistema podría también ser la calmodulina.

Con objeto de averiguar si la calmodulina estaba presente en el intestino delgado de conejo aplicamos el método utilizado por Debman para aislar la calmodulina de otros tejidos. Mediante el uso de la técnica de diálisis, medimos, en condiciones de equilibrio, la unión de la estelazina a la proteína aislada de intestino delgado de conejo en presencia de diferentes concentraciones de calcio. La concentración de calcio a la que se consiguió la mitad del valor máximo de la unión de estelazina a la proteína fue 10^{-6} molar, coherente con el valor obtenido para la constante de disociación de la estelazina cuando se medía su entrada en el tejido estimulado por la teofilina o el A23187. Asimismo, utilizando la técnica arriba mencionada, cuando se midió la unión de Ca^{2+}-45 a la proteína, a lo largo de una amplia gama de concentraciones de calcio en el medio de diálisis, la constante de disociación obtenida para el calcio fue próxima a 10^{-6} molar. Estos valores coincidían con los conseguidos por otros autores (Levin y Weiss) a propósito de la calmodulina aislada de otros tejidos, lo que inducía a pensar que la calmodulina también estaría presente en el intestino delgado del conejo. Así mismo dichos valores corroboraron la hipótesis de que los secretagogos ensayados aumentaban la concentración citosólica de Ca^{2+} hasta valores de al menos 10^{-6} molar. La RMI 12330A, al igual que inhibió la entrada de la estelazina al tejido estimulado por los secretagogos, inhibió en presencia de calcio la unión de la estelazina a la proteína aislada de íleon de conejo, lo que revelaba que la RMI 12330A actuaba sobre el complejo CDR-Ca.

Levin y Weiss descubrieron una buena correlación entre la capacidad de un determinado tranquilizante por unirse al complejo Ca-CDR, su acción inhibidora sobre la fosfodiesterasa y su efecto clínico como agente tranquilizante. En nuestro sistema secretor también observamos una buena correlación entre la unión de tres tranquilizantes (estelazina, pimozida y diazepan) al complejo Ca-CDR (CDR aislado de intestino de conejo) y su actividad antisecretora. La pimozida era el antisecretagogo más eficaz y tenía mayor afinidad por Ca-CDR que la estelazina, y ésta, a su vez, más que el diazepan. Esta gradación coincidía con la registrada por Levin y Weiss en su trabajo sobre la acción de estos tranquilizantes en la actividad de fosfodiesterasa. Cabía pensar, pues, que la acción antisecretora de estos fármacos se ejerce a través del complejo Ca-CDR.

A partir de estos resultados propusimos el siguiente modelo para explicar el mecanismo bioquímico intracelular que se modifica en respuesta a los secretagogos. La interacción de los secretagogos con las células del epitelio del intestino delgado conduce a un destacado aumento de la actividad citoplasmática de Ca^{2+}. Algunos secretagogos, tales como A23187, elevan la concentración intracelular de Ca^{2+} aumentando la permeabilidad al Ca^{2+} de la membrana de las células. Otros secretagogos (teofilina, VIP y toxina del cólera) incrementarían la concentración intracelular de AMPc, que, a su vez, movilizaría calcio de reservorios intracelulares, en donde normalmente se encontraría secuestrado. Un aumento en la concentración intracelular de calcio ionizado conduciría a la formación del complejo Ca-CDR, el cual actuaría sobre el borde en cepillo, elevando la permeabilidad de éste al cloruro. Tranquilizantes y RMI 12330A inhibirían la secreción intestinal uniéndose al complejo Ca-CDR y bloqueando su acción sobre la membrana.

Quedaría por averiguar el mecanismo por el cual el complejo Ca-CDR abre canales para el Cl^- a través del borde en cepillo. Podría actuar interaccionando directamente sobre alguna proteína de membrana (efector) o sobre alguna kinasa, que a su vez influyera sobre el efector.

En conclusión, los fundamentos experimentales descritos en este trabajo apoyan la hipótesis de que un mismo tipo celular es capaz de absorber y segregar agua y electrolitos. El balance entre estos dos procesos dependería de la concentración citoplasmática de Ca^{2+}. Cuando la concentración de $[Ca^{2+}]$ es baja se produce absorción neta, mientras que a altas concentraciones tiene lugar la secreción. La dirección y movimiento de agua y electrolitos está controlada por la permeabilidad de la membrana plasmática al Cl^-. La calmodulina, que actúa de transductor de señales intracelulares en un gran número de actividades celulares, también controla el movimiento y dirección del transporte neto de agua y electrolitos a través del intestino delgado. El AMPc, al igual que en otros tipos celulares, parece reforzar la señal del Ca^{2+} al aumentar su concentración citosólica.

La reproducción cooperativa de la abubilla arbórea verde

En estas aves de Africa oriental se reproduce una pareja de cada bandada. Los adultos no reproductores ayudan a la crianza de los polluelos. ¿Cómo se desarrolló esta pauta de comportamiento animal aparentemente altruista?

J. David Ligon y Sandra H. Ligon

Estos últimos años, los investigadores del comportamiento animal han dedicado especial atención a la evolución de organizaciones sociales complejas. Un aspecto del comportamiento social es la cooperación, es decir, el intercambio entre individuos de algún tipo de recurso o ayuda. La cooperación es común en las sociedades humanas y, por lo general, es mutuamente beneficiosa. Sin embargo, la prestación puede estar descompensada, ser incluso desinteresada (como en el heroísmo humano con sacrificio de la propia vida). En los insectos sociales se encuentran los ejemplos más comunes de intercambio desigual entre los animales inferiores. En estas sociedades animales las obreras son estériles y pasan toda su vida adulta dedicadas a actividades que benefician a la reproducción de la reina. A menudo dan su vida por defender el nido. Este comportamiento parece ir más allá de los límites de la cooperación y cuadra con la definición que el biólogo da del altruismo: "comportamiento que disminuye en potencia o realmente el rendimiento reproductor de la vida del altruista para beneficio de otro individuo o individuos". Naturalmente, esta definición no implica ningún razonamiento consciente o premeditación.

Aunque los insectos sociales proporcionan los ejemplos más sorprendentes de altruismo aparente, también se encuentran niveles de cooperación complejos, que incluyen la prestación de ayuda manifiestamente desigual, en los animales superiores, aparte del hombre. Un vívido caso en cuestión es el de una especie de ave insectívora de Africa: la abubilla arbórea verde (*Phoeniculus purpureus*). La unidad social de estas aves, la bandada, puede tener hasta 16 miembros, pero sólo una pareja reproductora. Los demás miembros de la bandada sexualmente maduros pero no reproductores sirven a la vez de "asistentes de nido", compartiendo la carga de acarrear alimento para la hembra que incuba y después para los polluelos, y de "guardas", defendiendo de los depredadores a los pollos, volantones luego, y participando durante todo el año en actividades propias de la bandada, como la protección del territorio.

Más de 100 especies de aves comparten con la abubilla arbórea verde esta pauta de reproducción cooperativa, pero no hay ningún elenco cerrado de factores que expliquen por qué una especie exhibe la pauta de comportamiento y otra especie aparentemente similar no lo hace. Para comprender las presiones de selección que favorecen la evolución de la reproducción cooperativa debe tenerse en cuenta la influencia del ambiente sobre un sistema social determinado. Entre los factores ambientales están el clima, la distribución del alimento tanto en el tiempo (con las estaciones) como en el espacio (dentro de la zona utilizada), la disponibilidad y calidad de los lugares para el descanso y para establecer el nido y el tipo y número de depredadores. Otro componente importante de la selección natural es lo que puede denominarse ambiente social: el efecto sobre el individuo de su interacción con otros miembros de la misma especie.

Iniciamos nuestro trabajo con las abubillas arbóreas en 1975. Estos fenicúlidos constituyen un sujeto de estudio muy idóneo. Las bandadas que observamos viven en bosques de acacias cerca del lago Naivasha, en Kenia, donde por la noche descansan en el interior de cavidades de los troncos de las acacias, algunas de ellas formadas de modo natural y, otras, agujeros desocupados de picos carpinteros. Lo que facilita la captura y marcado de los individuos; basta con tapar la boca de la cavidad al oscurecer y, a la mañana siguiente, se espera a que el ave, al salir, se introduzca en una bolsa transparente. La claridad del bosque y el escaso follaje de las acacias ayudan también a la observación de las aves. Asimismo facilita la tarea su conducta, por lo general indiferente a la presencia de seres humanos, y su tamaño, relativamente grande (los machos suelen tener unos 36 centímetros de longitud y, las hembras, 30).

Durante nuestra primera campaña, desde julio de 1975 hasta mayo de 1976, pusimos bandas de color a 151 aves, total que incluía a casi todos los miembros de 25 bandadas. Posteriormente marcamos a otras 218 aves, la mayoría de ellas descendientes de las primeras bandadas que habíamos marcado. En nuestra libreta de campo anotábamos el sexo, la edad y la afiliación a las distintas bandadas de las aves, así como la ascendencia de casi todos los componentes del segundo grupo.

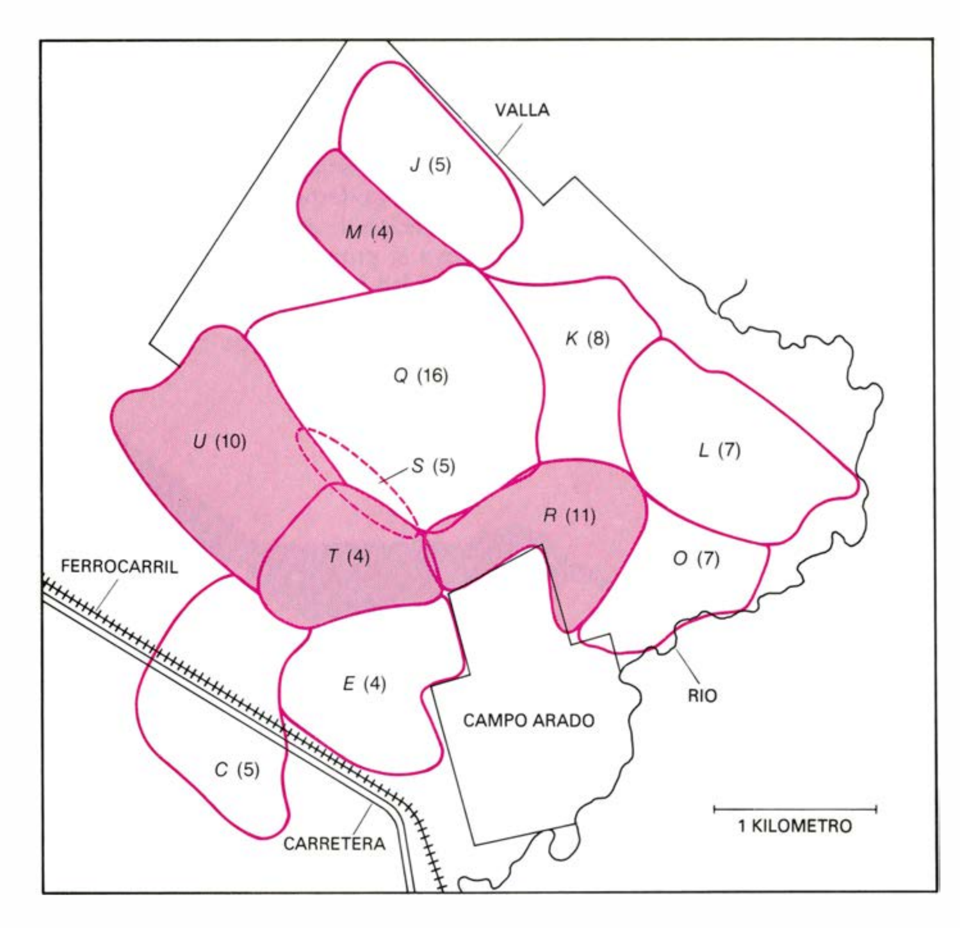
Las aves de nuestra zona de estudio eran muy territoriales. Era insólito que un hábitat adecuado, por ejemplo un bosquete de árboles que tuvieran cavidades que sirvieran de descansaderos, no se lo "apropiara" una u otra bandada. Los territorios se defendían vigorosamente contra la intrusión de aves procedentes de bandadas vecinas y contra las abubillas arbóreas "forasteras" que ocasionalmente vagaban por la zona. Esta pauta de comportamiento hacía raro el establecimiento de nuevos territorios. Los grupos diferían en razón del tamaño; algunos constaban exclusivamente de la pareja reproductora. Mas, por lo general, las bandadas eran mayores, y las aves que "ayudaban" a la pareja reproductora solían ser, aunque no invariablemente, los hermanos de una de las aves reproductoras o los hijos de una de las aves de pareja reproductora, o de ambas.

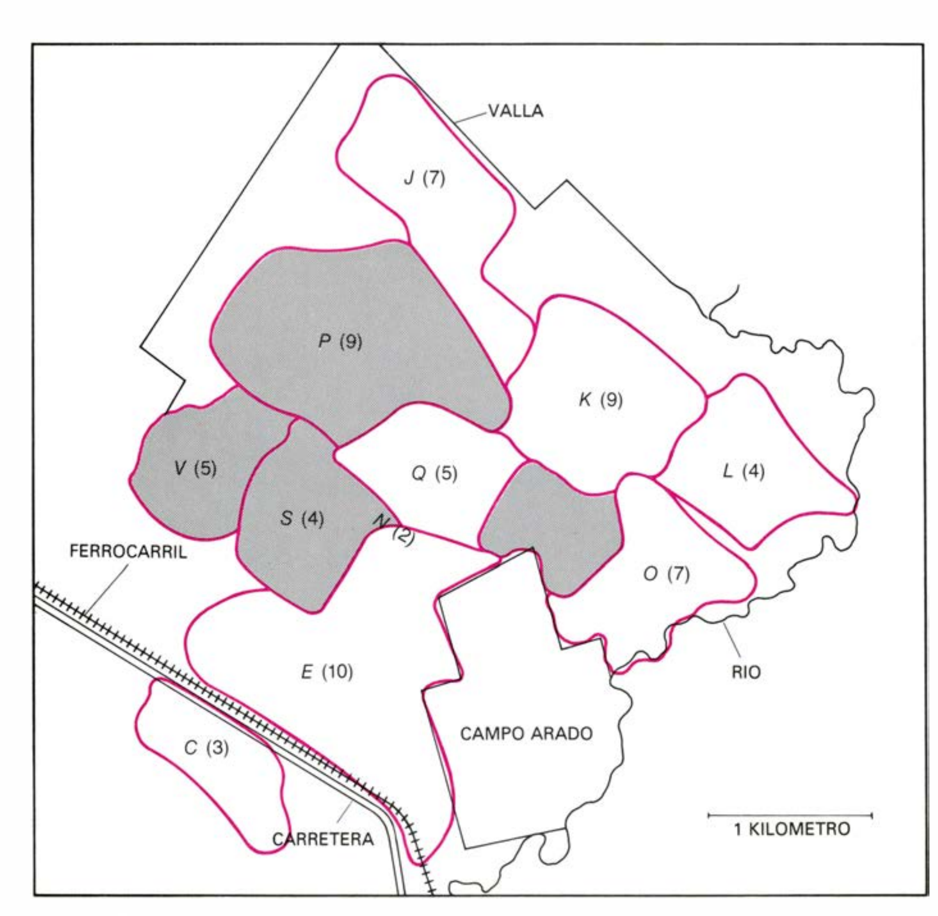
Las abubillas arbóreas de nuestra zona de estudio no vivían mucho, por término medio. Cada año moría entre el 30 y el 40 por ciento de la población. La distribución de la tasa de mortalidad era desigual. Las bajas más frecuentes correspondían a los machos reproductores; el número de muertos en este grupo superaba al de las tres categorías restantes: hembras reproductoras, hembras asistentes y machos asistentes. La mayoría de las muertes debíanse a depredación. Los principales depredadores resultaron ser jinetas (carnívoros parecidos a gatos de la familia de las mangostas) y hormigas cazadoras (el equivalente africano de las hormigas legionarias del Nuevo Mundo). Ambos tipos de depredadores realizan incursiones nocturnas en los descansaderos de las abubillas arbóreas. En varias ocasiones encontramos los restos de aves muertas, todavía en los agujeros de descanso o en el suelo bajo los puntos de anidación.

Además de la depredación, hay otros dos factores ambientales que dejan sentir su poderosa influencia sobre la población de abubillas arbóreas. Se refiere el primero a la oportunidad y la cantidad de precipitación. En esta región de Kenia las lluvias son muy variables. La pauta anual "normal" com-



"ASISTENTES" NO REPRODUCTORES de una bandada de abubillas arbóreas verdes rodean a un polluelo (*centro*) en esta representación artística. Dos de ellos ofrecen a la abubilla arbórea juvenil el principal alimento local, orugas, mientras otro lo está aseando. Los tres asistentes restantes, con sus picos abiertos, vocalizan en dirección al polluelo. Una bandada de abubillas arbóreas puede contar con hasta 16 miembros, pero sólo una pareja adulta se reproduce. Los adultos no reproductores de la bandada toman parte en la crianza de los descendientes y ayudan asimismo a la defensa del territorio.





BOSQUE CLARO DE ACACIAS, en una gran finca del valle del Rift, Kenia, que fue la zona seleccionada para el estudio de las abubillas arbóreas. A finales de 1975 (*arriba*), este terreno de 800 hectáreas estaba habitado por 11 bandadas establecidas, y una duodécima se hallaba en trance de establecerse (*línea discontinua a la izquierda del centro*), Se indica el número de aves por cada bandada al finalizar el año. El total era de 86. Los cuatro territorios en color habían desaparecido a finales de 1981 (*abajo*), pero se habían establecido cuatro nuevos territorios, señalados en gris. El número de aves era entonces de 65.

prende una estación seca que se extiende desde diciembre hasta febrero, seguida por las "lluvias largas", que empiezan en marzo o abril. A lo largo de los siete años de nuestro estudio las abubillas arbóreas iniciaron su actividad reproductora en mayo o junio, unas pocas semanas después del comienzo de las lluvias largas.

La pluviometría anormal puede limitar drásticamente la principal fuente de alimento de las abubillas arbóreas: las larvas de unas 2000 especies de polillas que viven en el bosque. Si la estación seca es verdaderamente seca, las polillas que se hallan pupando prosperan en sus capullos enterrados. Cuando comienzan las lluvias, las polillas salen para reproducirse y poner los huevos que darán origen a la siguiente generación. Pero si llueve durante la estación seca, las pupas se pudrirán o las destruirán hongos y otros microorganismos. Al reducirse el número de crisálidas que han de convertirse en polillas y poner huevos, habrá menos orugas.

En 1979 fuimos testigos de una situación desfavorable de ese tipo. Aquel año cayeron cerca de 30 centímetros de agua de lluvia durante los meses de enero y febrero, cantidad que hemos de ponderar en relación con una media de menos de seis centímetros a lo largo de 34 años. Las abubillas arbóreas comenzaron a anidar en junio, pero sólo dos de las 11 bandadas de nuestra zona de estudio consiguieron criar algún polluelo. En 1981 ocurrió exactamente lo contrario. En enero y febrero cayeron 2,1 centímetros de lluvia, y las crisálidas de las polillas prosperaron. Las lluvias largas de marzo a mayo fueron extraordinariamente copiosas: 36 centímetros. Las 19 bandadas de abubillas arbóreas que aquel año se asentaban en nuestra zona, entre las que había cuatro parejas sin asistentes, anidaron en mayo y junio. El éxito reproductor fue insólitamente elevado, y todos los grupos, salvo dos, produjeron descendientes. Todavía hubiera sido mayor si los despredadores no hubieran devorado dos polladas. Esta gran variación anual en el suministro de alimento (y en el éxito reproductor) representa una variante ambiental crítica para las abubillas arbóreas.

Otra variable, quizás igualmente decisiva, la representa el número de agujeros de que disponen las abubillas arbóreas para descansar. Compiten por las cavidades con aves de otras especies y también con abejas melíferas y con pequeños mamíferos. La presencia o ausencia de agujeros-descansaderos en una zona de bosque determina la distri-

bución de las bandadas de fenicúlidos y el éxito reproductor de las mismas a lo largo de los años. Los descansaderos adecuados suelen ser tan escasos que la mayoría de los miembros de cada sexo, si no todos, reposan juntos. Este comportamiento de "dormitorio comunal" puede ser un subproducto del dimorfismo sexual: la diferencia de tamaño entre los machos y las hembras. Las hembras, que son más pequeñas, pueden penetrar en cavidades donde los machos, mayores, no pueden hacerlo. En territorios donde los agujeros adecuados para descansar eran raros encontramos hasta ocho hembras descansando juntas en una cavidad.

Esta segregación por el sexo tiene una importante implicación territorial. Un depredador nocturno puede, por casualidad, eliminar a la mayor parte o a todos los machos o hembras de una bandada en una noche. Un tal acontecimiento deja el territorio de la bandada abierto a la colonización por las aves del sexo que ha sido eliminado; los colonizadores pueden ser individuos solitarios o miembros de bandadas de territorios adyacentes.

A lo largo de nuestros años de observación, fuimos anotando numerosos ejemplos del significado adaptativo y evolutivo del comportamiento de las abubillas arbóreas asistentes. Considérese, en efecto, la elevada tasa de mortalidad de las aves. Para un pollo volantón será muy arriesgado abandonar el territorio de sus padres y buscar un territorio vacío en otro lugar sin un descansadero conocido en el que cobijarse por la noche. También lo será, indirectamente, para los padres: en un sentido evolutivo, producir descendientes maduros es la única razón de su existencia. Para el volantón, permanecer en el territorio de sus padres desempeñando el papel de un asistente no reproductor es mucho menos arriesgado. Puesto que los asistentes suelen ser los descendientes de uno de los miembros de la pareja reproductora de la bandada (o de ambos), permanecer con ésta durante un período temporal suplementario beneficia asimismo en un sentido evolutivo a uno de los padres, o a ambos. Efectivamente, consideramos que esta pauta de afiliación continuada constituye una forma de cuidado paterno prolongado.

Un hecho observado que apoya esta interpretación de la afiliación continuada por parte de las aves asistentes es que no hay correlación positiva entre el número de asistentes de una bandada y el número de descendientes producidos en una estación por la pareja reproductora. Ello significa que la protección que la afiliación continuada a la bandada proporciona a los asistentes no aumenta necesariamente el rendimiento reproductor de la pareja que cría, y a veces puede incluso reducirlo.

Un aspecto ulterior de la elevada mortalidad es la probabilidad de que uno u otro de los componentes de la pareja reproductora de una bandada muera y sea sustituido en un plazo breve. Valga de ejemplo que, en 1981, un año de éxito de cría excepcionalmente elevado, las 12 bandadas en nuestra zona de estudio comprendían únicamente 25 asistentes. De este total, cinco asistentes (el 20 por ciento) tenían los mismos padres que los pollos a los que los asistentes estaban alimentando. El grado de parentesco entre los otros 20 asistentes y los pollos que cuidaban oscila-



DISPUTA TERRITORIAL, dibujada a partir de las observaciones de campo. Tres aves de una bandada se enfrentan a cuatro de una bandada vecina. La disputa es principalmente vocal, sin embargo, un ave de cada grupo permanece con las alas semiabiertas. (La ilustración es de Patricia J. Wynne.)

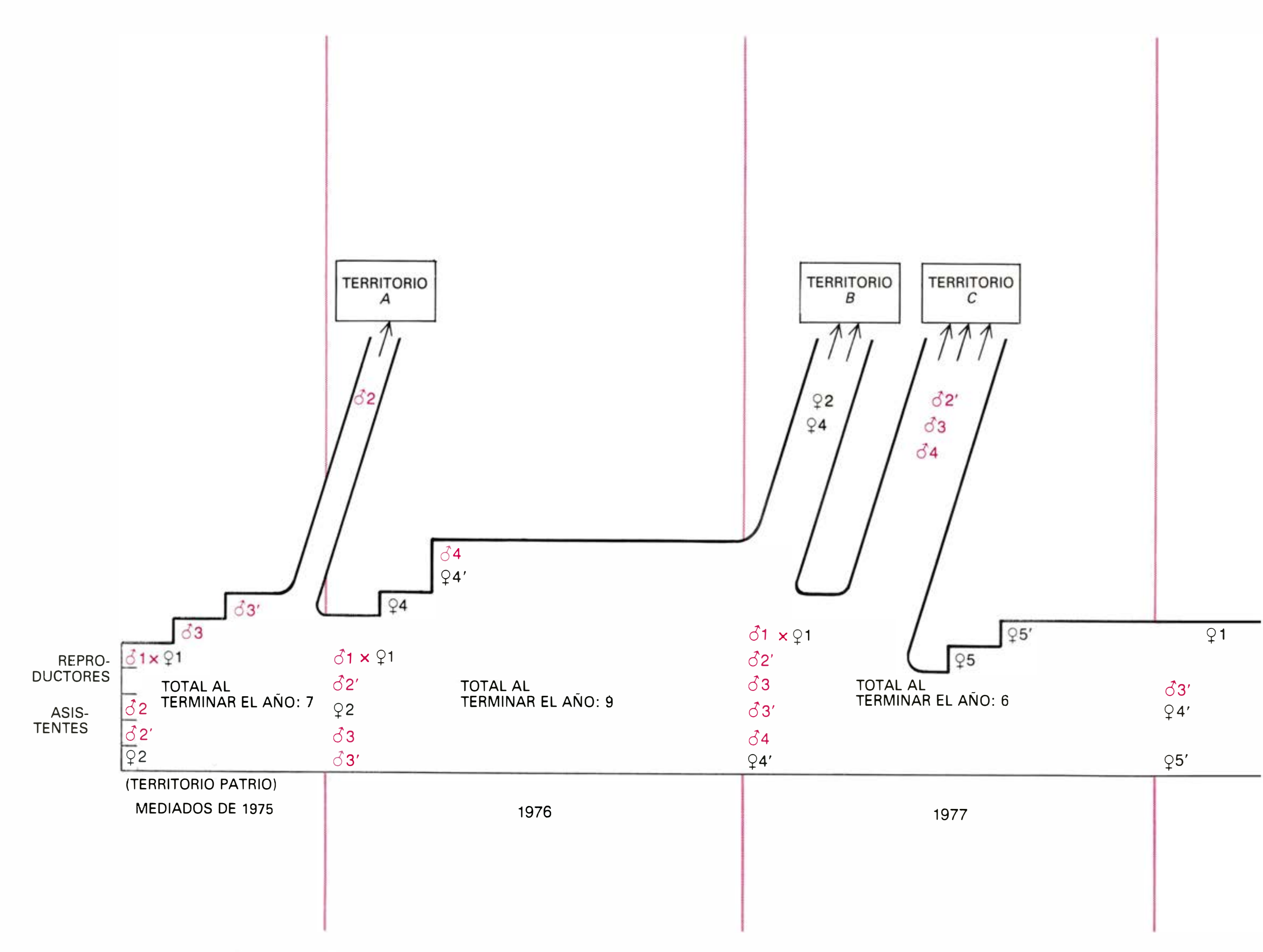
ba del cero (tres aves de las que se sabía que no compartían parientes cercanos) al 37,5 por ciento (seis aves con un progenitor en común con los pollos y los otros progenitores hermanos). En otras palabras, la mortalidad adulta elevada conducía a niveles de parentesco entre los asistentes y los pollos más bajos de los que podría predecir la teoría de parentesco.

Los asistentes no suelen tener la opción de colonizar un nuevo territorio, al menos durante la mayor parte del tiempo. Por ejemplo, el número de abubillas arbóreas de nuestra zona de estudio principal fluctuó ampliamente entre agosto de 1975 y agosto de 1981, variando desde un mínimo de 46 aves hasta un máximo de 94. Sin embargo, el número de territorios que contenían una pareja reproductora al inicio de cada estación de cría fluctuó mucho menos, desde un mínimo de 12 territorios hasta un máximo de 14. Entre 1975 y 1981 cuatro bandadas desaparecieron; sus antiguos territorios quedaron vacíos. En el mismo período se desarrollaron cuatro nuevos territorios.

Si se considera la gran variación en el número de individuos a lo largo del intervalo de siete años, la relativa estabilidad en el número de reproductores y de territorios sugiere que la oportunidad de colonizar un nuevo territorio es rara. Evidentemente, una pauta de vagabundeo (es decir, de búsqueda efectiva de un nuevo territorio), con los peligros que la acompañan, es insólita por la sencilla razón de que, en general, resulta desventajosa. Estas consideraciones ayudan a explicar por qué las jóvenes abubillas arbóreas permanecen con la bandada de sus padres. Pero no justifican el comportamiento asistencial de las aves.

Esto no significa que la emigración constituya siempre una desventaja. La elevada tasa de mortalidad de las abubillas arbóreas adultas revela que los territorios vecinos tienen frecuentes vacantes reproductoras. La manera más común de que un ave madura no reproductora de uno u otro sexo alcance la categoría de reproductor es emigrando hacia un territorio donde se ha creado un vacío por la muerte de un reproductor del mismo sexo (o de los no reproductores maduros del mismo sexo, si hay alguno).

Raramente emprenden los emigrantes estos cambios en solitario. Observamos que un desplazamiento con éxito a



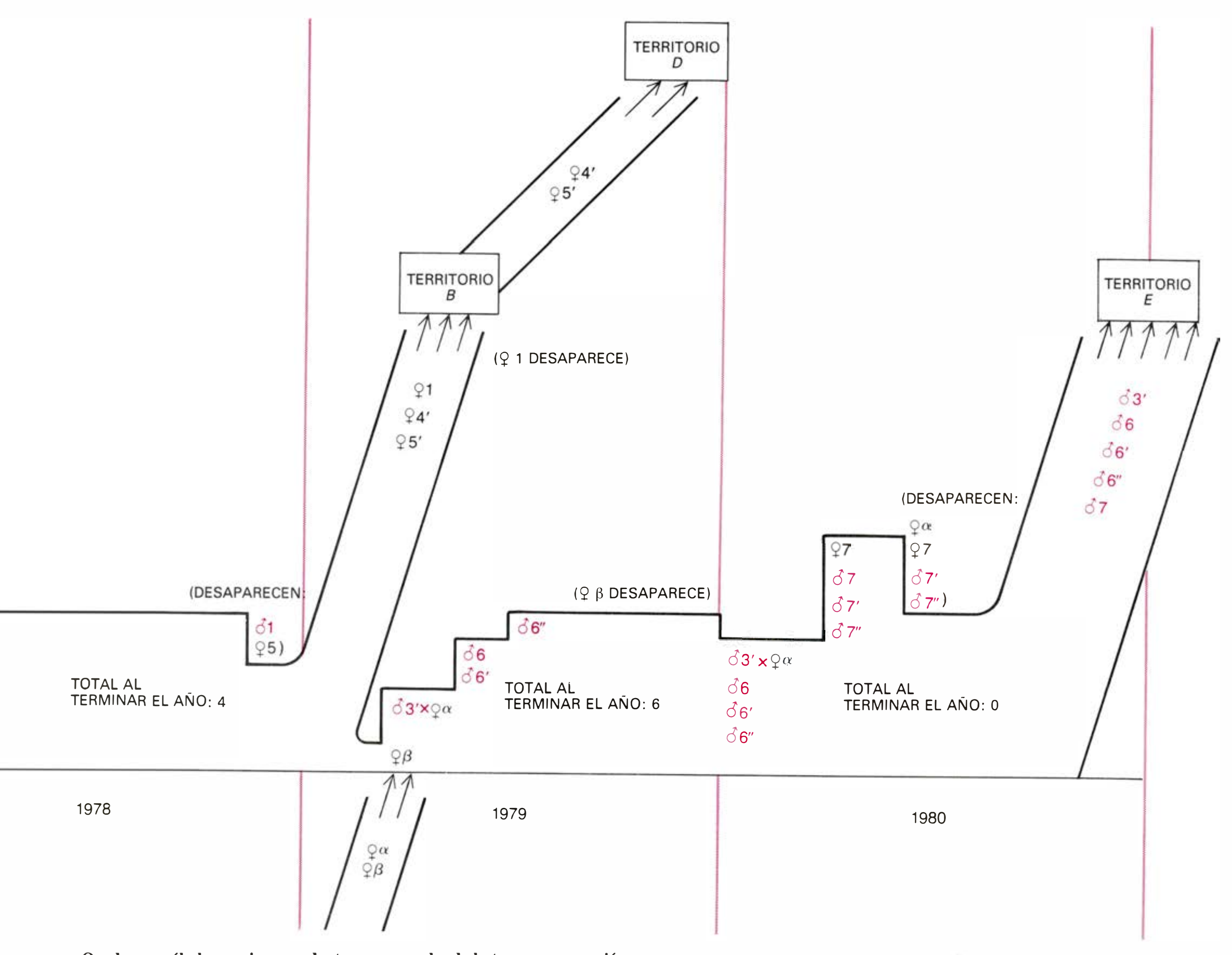
SECUENCIA DE CAMBIO en una bandada de abubillas arbóreas a lo largo de un intervalo de 66 meses, que se indica en este esquema que recoge el éxito reproductor, las emigraciones, las inmigraciones y las desapariciones de cada ave. Al marcar las aves por primera vez, en un territorio adyacente a la zona de estudio principal, la bandada comprendía una pareja reproductora (*macho 1 y hembra 1; los números los identifican como la primera generación observada*) y otras tres aves maduras de una segunda generación (*machos 2 y 2′ y hembra 2*), todos ellos probablemente descendientes de la pareja reproductora. En la estación reproductora de 1975 se añadieron a la bandada otras dos aves (*machos 3 y 3′*). A principios de 1976, el macho 2 emigró solo al territorio *A*. En la estación reproductora de 1976 se añadieron a la bandada tres descendientes (*hembra 4 y machos 4 y 4′*). A principios de 1977 dos aves (*hembra 2 y hembra 4*) emigraron al territorio *B*, de acuerdo con la pauta normal de emigración: un ave más vieja se acompaña de un subordinado más joven del mismo sexo. Más avanzado 1977, otros tres emigrantes abandonaron la bandada para dirigirse al territorio *C* (*macho 2′, macho 3 y macho 4*).

un territorio extraño solía implicar a un grupo de dos o más aves del mismo sexo. Estos grupos estaban compuestos generalmente por compañeros de bandada, pero nunca incluían compañeros de nido; el antagonismo entre antiguos compañeros de nido es fuerte y no emigran juntos. En consecuencia, los grupos emigrantes constaban de un no reproductor más viejo (los llamábamos aves alfa) y uno o dos más jóvenes (aves beta y gamma). Era evidente una sutil jerarquía de dominancia, correlacionada positivamente con la mayor edad. Observamos que el emigrante alfa se convertía en el reproductor sustituto en la bandada de adopción. Si el emigrante alfa moría después, el emigrante beta heredaba la categoría reproductora del alfa.

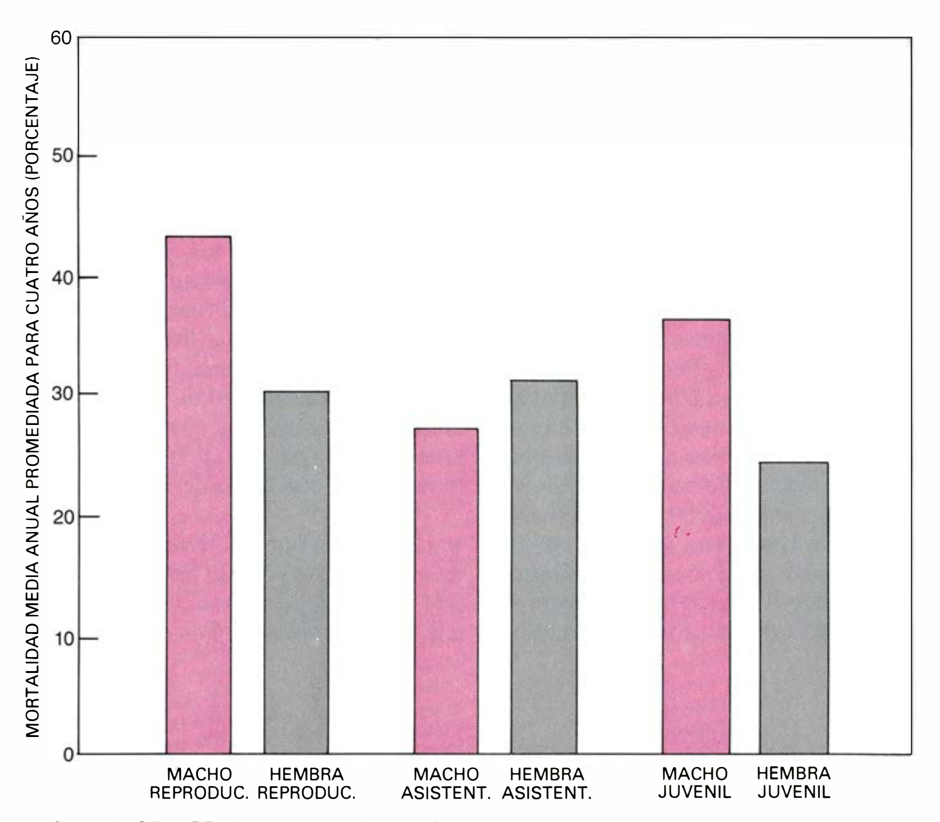
Los hechos relativos a la emigración de grupo, en la que el socio más viejo dominante era un asistente que había participado en la cría del socio (o socios) subordinado más joven, nos proporcionaron un atisbo de la evolución del comportamiento asistencial. Para que un no reproductor accediera a la categoría de reproductor por emigración (uno de los dos caminos hacia la supervivencia en el sentido genético), debía primero adquirir uno o más aliados subordinados. ¿Podría ésta ser, al menos en parte, la razón por la que los asistentes prestan ayuda? Nada hay en este proceso evolutivo que requiera implicar el ejercicio del raciocinio: los genes de los no reproductores que no ayudan y, por tanto, que no están asociados, corren el riesgo de quedar eliminados del acervo génico de la especie en el transcurso de una sola generación.

¿Cómo se establece el vínculo entre el asistente y el polluelo al que ayuda? Descubrimos que los asistentes no sólo llevaban alimento a los pollos, sino que también robaban orugas de otros asistentes con el propósito de dárselas a los pollos a los que asistían. Los asistentes también atendían a los polluelos de otras maneras: los aseaban, se posaban cerca y vocalizaban dirigiéndose hacia ellos. Ninguna de estas actividades, ni siquiera ir a buscar alimento, parece representar una inversión costosa en tiempo y energía si se compara con la gran ganancia genética a la que se aspira: lograr la condición de reproductor.

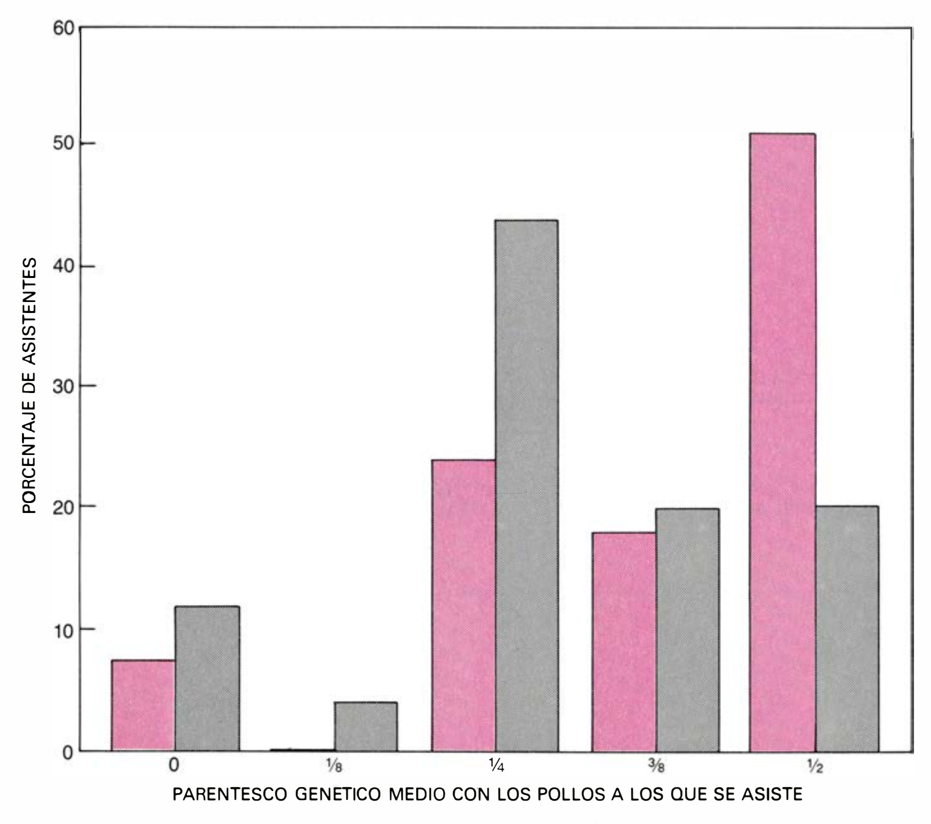
El hecho de que los no reproductores más viejos puedan beneficiarse de este modo deja todavía sin respuesta una pregunta clave: ¿Por qué las aves



Quedan ya sólo la pareja reproductora, un macho de la tercera generación y una hembra de la cuarta generación. A finales de año, la pareja reproductora anidó dos veces, criando dos hembras de la quinta generación. Esa fue la última estación para el macho 1; desapareció en 1978, al igual que una de las hembras jóvenes, la 5. Ese año no hubo descendencia. A mediados de 1979, la hembra reproductora, la hembra 4′ y la hembra 5′ emigraron al territorio *B*, dejando al macho 3′ como ocupante único del territorio patrio. La madre emigrante desapareció pronto, y las dos hembras más jóvenes se desplazaron al territorio *D*. Por aquella época dos hembras emigrantes de una bandada adyacente (*alfa y beta*) se unieron al macho 3′. Al finalizar el año, el macho 3′ y la hembra alfa se convirtieron en la segunda pareja reproductora en el territorio patrio. Produjeron la sexta generación: los machos 6, 6′ y 6″. En 1980 dieron otros cuatro descendientes: la hembra 7 y los machos 7, 7′ y 7″. Tanto la hembra alfa como la hembra beta, la hembra 7, el macho 7′ y el macho 7″ desaparecieron aquel año. El macho reproductor y sus descendientes (machos 6, 6′, 6″ y 7) emigraron al territorio *E*, dejando vacío el patrio.



TASA DE MORTALIDAD de las abubillas arbóreas a lo largo de un período de cuatro años (de enero de 1976 a diciembre de 1979); aunque alta, no se distribuye de manera uniforme, ni en cuanto al sexo ni en cuanto a la condición social. Los machos reproductores y los machos juveniles encabezan las bajas; las hembras juveniles y los machos asistentes se sitúan en la cola. Las tasas correspondientes a las hembras reproductoras y a las hembras asistentes fueron muy parecidas, y se sitúan entre las de las otras aves.



VINCULOS DE PARENTESCO dentro de cada bandada; se atenúan con el tiempo. Este gráfico compara la relación entre los asistentes y los polluelos que cuidaron en 1978 (*color*) y 1981 (*gris*). El primer año había 14 bandadas, con un total de 55 asistentes; el último, nueve bandadas, con 25 asistentes en total. En 1978, más del 50 por ciento de los asistentes compartían una ascendencia común con los polluelos. En 1981, sólo el 20 por ciento de los asistentes se hallaba emparentado con los pollos. Menos del 25 % de los asistentes, en 1978, guardaban con los polluelos una relación de parentesco de medio hermanos, y ninguno equivalía a un primo hermano. En 1981, esos parentescos se acercaban al 45 y 5 %, respectivamente.

más jóvenes acompañan a sus mayores cuando surge la oportunidad de emigrar? Los beta no sólo ayudan a los alfas a establecerse en un nuevo territorio, sino que cargan con las obligaciones de un asistente cuando eclosionan los hijos de los alfa. Aquí también, la respuesta parece residir en la pauta de mortalidad de las abubillas arbóreas. Sin que ello suponga comportamiento racional alguno, una descripción antropomórfica de las acciones de las aves jóvenes no sería que éstas "devuelven el favor" a las aves más viejas por su ayuda anterior, sino más bien que "juegan con ventaja", a la espera de su propia recompensa genética.

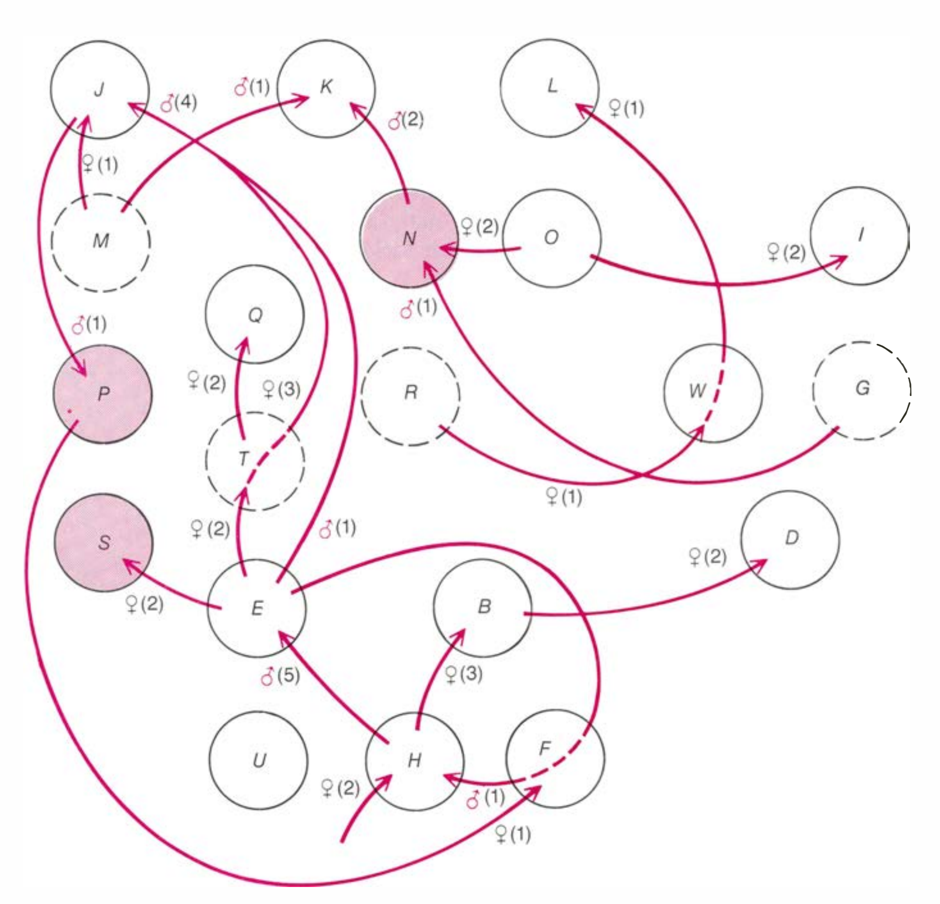
Por ejemplo, recuérdese que los machos alfa tienen una tasa de mortalidad significativamente superior a la de los machos beta. En consecuencia, los machos subordinados sobrevivirán frecuentemente a sus socios dominantes y heredarán la condición de reproductores. Observamos que en grupos de dos a cuatro machos emigrantes, y en bandadas recién formadas compuestas de sólo dos machos adultos y una o dos hembras adultas, en 10 de 13 casos los machos alfa murieron antes que los machos más jóvenes. También descubrimos que a lo largo de su vida los machos emigrantes que en un tiempo habían sido subordinados dejaban tantos descendientes supervivientes como los machos originalmente dominantes. Esta observación guarda una relación estricta con el concepto evolutivo conocido por eficacia individual. En lo que se refiere al comportamiento de las aves beta, así como en lo que concierne al de las aves alfa, no hay razón para suponer que, cuando emigran con compañeros de bandada más viejos, estén, en modo alguno, actuando en contra de sus intereses reproductores.

Vimos asimismo que algunas hembras asistentes, a diferencia de los machos asistentes, abandonaban voluntariamente su territorio original sin la compañía de subordinados y erraban por una zona extensa, aparentemente en busca de un territorio vacío. Este comportamiento puede surgir del hecho de que la tasa de mortalidad de las hembras reproductoras es significativamente menor que la de los machos reproductores. Por tanto, la probabilidad de que una hembra asistente consiga la condición de reproductora a una edad temprana en la bandada propia es remota, y buscar la condición de reproductora en algún otro lugar, aunque sea arriesgado para un ave solitaria, puede compensar desde el punto de vista genético.

Puesto que el futuro de cualquier ave en particular es incierto, resulta genéticamente ventajoso para las abubillas arbóreas producir crías que muestren hábitos de comportamiento distintos. Ilustraremos este punto con la descripción de las observaciones iniciadas en 1977. Ese año, tres de cuatro machos asistentes hermanos, pertenecientes a una bandada compuesta por la pareja reproductora usual y cinco no reproductores, emigraron a un territorio vecino. Allí expulsaron al macho único que lo ocupaba y se unieron a tres hembras. El cuarto macho asistente no emigró. Dos años más tarde, el macho padre murió y la madre del cuarto macho asistente y dos hermanas emigraron a continuación, dejando al macho asistente heredero único del territorio.

Este macho pronto se apareó con una de dos hembras emigrantes procedentes de una bandada vecina. A finales de 1979 y en 1980 la nueva pareja reproductora crió un total de siete descendientes. En contraste con este récord de éxito reproductor, dos de los tres machos hermanos emigrantes habían muerto para finales de 1979 sin haber dejado descendientes. El más joven de los tres vivía todavía a primeros de 1982, pero sólo había producido un volantón. Si el cuarto hermano no se hubiera quedado en casa, su contribución al acervo génico de los fenicúlidos es probable que hubiera sido nula.

Aunque muchos grupos emigrantes estaban formados por hermanos, observamos también que había abubillas arbóreas no emparentadas del mismo sexo que se asociaban para fundar el núcleo de una nueva bandada. Ello ocurría, por norma, cuando un ave más vieja y dominante permitía que un individuo más joven y subordinado se uniera a ella. Observamos este comportamiento sólo cuando un macho más viejo necesitaba un aliado para defender y mantener su territorio. En ocasiones, dos hembras permitían que una tercera no emparentada se les uniera en las mismas circunstancias. De los 40 casos observados de formación de una nueva bandada (o de la sustitución de todos los miembros de un sexo, o de todos los miembros de un sexo menos uno), 17 de los grupos contenían no parientes del mismo sexo y otros 17 contenían parientes seguros o probables del mismo sexo. En seis casos, los lazos de parentesco nos eran desconocidos.



MOVIMIENTOS DE LOS EMIGRANTES entre 20 territorios de abubillas arbóreas durante un período de cuatro años (1978-81). Allí donde los territorios coinciden con los de las ilustraciones de las páginas 86 y 88-89, las letras que los identifican se corresponden. Se muestran 31 migraciones individuales, que corresponden a los vuelos de 14 machos y de 17 hembras. Las aves que penetran en un territorio y luego lo abandonan para dirigirse a otro (por ejemplo, de *E* a *J* pasando por *T*) se contabilizan sólo una vez. Sin embargo, las mismas aves pueden desplazarse más de una vez. Por ejemplo, dos de las tres hembras que pasaron de *H* a *B* se desplazaron a continuación de *B* a *D*. Los cuatro territorios cuyo perímetro se indica a trazos fueron abandonados, y los tres señalados en color se colonizaron durante este período.

Permítasenos ahora sugerir de qué modo pudo haberse originado el comportamiento reproductor cooperativo en las abubillas arbóreas verdes y, a su vez, cómo este comportamiento condujo a la ayuda mutua entre individuos no emparentados del mismo sexo. En primer lugar, como hemos visto, hay tres características del ambiente y del ciclo biológico de los fenicúlidos que parecen haber dispuesto el escenario para el desarrollo de la reproducción cooperativa, al primar a las parejas reproductoras que permitían la estancia indefinida en el territorio paterno, al menos a algunos de sus descendientes. Las tres características son una elevada tasa de mortalidad de los adultos, una tasa de natalidad impredecible debido a un suministro alimentario fluctuante y la insegura disponibilidad de descansaderos.

El período de residencia prolongado para los descendientes adultos otorga un doble beneficio a la pareja reproductora. Consiste el primero en la oportunidad de expansión territorial de que dispone una bandada que está compuesta por más de dos adultos. En este ambiente uniforme de bosque de acacias, un territorio mayor significa una zona de alimentación más amplia. De manera similar, la presencia de más de dos aves adultas revela que la defensa del territorio recae con menos peso sobre cada uno de los adultos.

El segundo beneficio que arrastra un período de residencia prolongado yace en la ayuda, del tipo que sea, que los jóvenes adultos pueden ofrecer a la pareja reproductora en la alimentación y las protección de sus polluelos. La presencia de asistentes probablemente reduce el coste de la anidación anual para la pareja reproductora, con lo que aumenta el período potencial de reproducción de la pareja. Luego, cuando surge la oportunidad de ocupar nuevo territorio para los adultos no reproductores, estos emigrantes ven aumentada su probabilidad de éxito, ya que disponen de socios entre los compañeros de bandada más jóvenes del mismo sexo.

Una vez establecida entre individuos emparentados esta interdependencia para la defensa y la expansión territorial, así como para la reproducción y posteriormente para la adquisición de nuevos territorios, cualquier pareja de abubillas arbóreas, o cualquier ave solitaria, no asistidas, se hallan en situación de una abrumadora desventaja competitiva. Si la mayoría de estas aves

han de aduenarse de un territorio por ellas mismas y, eventualmente, reproducirse, la mejor opción que tienen una vez han encontrado pareja es buscar socios no emparentados de su mismo sexo. Y esto es lo que hacen al permitir que aves no afiliadas más jóvenes y subordinadas se les unan.

¿Hasta qué punto los hechos del sistema social de las abubillas arbóreas concuerdan con las diversas teorías que explican el comportamiento cooperativo? Tres teorías revisten particular interés: la selección darwiniana individual, la selección de parentesco y la reciprocidad. La teoría de la selección individual, según la expuso Charles Darwin, se basa en la premisa de que los animales se comportan de suerte tal que se maximicen las oportunidades de producir sus propios descendientes. En el contexto de la reproducción cooperativa en las aves, un asistente de nido de uno u otro sexo puede ganar, a la larga, recursos valiosos, como territorio, por ejemplo, si primero sirve como aprendiz de los reproductores, más viejos y socialmente dominantes.

TERRITORIOS	NUMERO DE MACHOS	ALFA	BETA	GAMMA	DELTA
H	3	(1)	(2)		
A	4	(1)		(1)	(1)
S		(1)	(4)	(3)	(2)
Q	2	(1)	(2)		
Q	2	(2)	(1)		
E	2	(1)	(2)		
M	2	(1)	(2)		
K	2	(1)	(2)		
L	2	(2)	(1)		
X	2	(2)	(1)		
O	2	(1)			
J	4	(1)	(2)		(1)
F	2	(1)	(2)		
	PROMEDIO	2.23	3.00	.33	

EXITO REPRODUCTOR de los machos emigrantes de abubillas arbóreas. Nueve parejas de machos y cuatro grupos de machos más numerosos emigraron desde los territorios que se identifican en el extremo izquiedo. Un círculo vacío indica falta de éxito reproductor para el ave en cuestión. (Alfa es el macho más viejo, beta el segundo menos viejo, y así sucesivamente.) Un círculo lleno indica el máximo éxito reproductor observado: 11 engendrados por el ave beta del territorio *E*. La porción de color de los demás círculos es proporcional al éxito reproductor de ese ave. El número que acompaña los círculos indica el orden de fallecimiento en el grupo; por ejemplo, de los cuatro emigrantes del territorio *A* tres murieron simultáneamente. El macho beta de este grupo todavía seguía vivo cuando concluyeron las observaciones. Aunque los alfa fueron los primeros en reproducirse, fue mayor el éxito reproductor medio de los beta.

La selección de parentesco, una extensión de la selección individual, es una teoría propuesta (en primer lugar por W. D. Hamilton, de la Universidad de Michigan en Ann Arbor) para explicar lo que parecía ser comportamiento altruista. Al objeto de dar cuenta del altruismo aparente que se observa en la naturaleza, Hamilton propuso que, si el comportamiento generoso que muestra un individuo se dirigiera hacia un pariente, serviría para promover los genes compartidos por el altruista y el receptor de la ayuda de éste, de modo que los costes y los beneficios de este comportamiento deberían estar correlacionados con el grado de parentesco genético entre los dos individuos que interactúan.

La idea subyacente a la tercera teoría, la de la reciprocidad, es la siguiente: un individuo proporciona ayuda a otro con la esperanza de que será compensado por ello (con un valor igual o superior al de la ayuda, pero no necesariamente en la misma forma). La reciprocidad es una manera especializada de obtener beneficios para el individuo, y por ello no difiere, en esencia, de la selección darwiniana individual.

Ahora bien, puesto que la mayoría de las aves que interactúan en el seno de la unidad social cerrada de las abubillas arbóreas son parientes, podría razonarse que lo que ha sido crucial para la evolución del sistema es la selección de parentesco: aquella selección natural en la que los animales ayudan de preferencia a los parientes, a costa quizá de sí mismos. Efectivamente, un lazo de parentesco ha sido claramente fundamental para la evolución del sistema social de los fenicúlidos: el vínculo entre el progenitor y los hijos. Sin embargo, no está claro si son necesarios otros lazos de parentesco para el mantenimiento del sistema. Por ejemplo, en tres o cuatro bandadas hemos observado cada año, al menos desde 1977, asistentes que no estaban emparentados con los polluelos a los que cuidaban. Algunos de estos no parientes, además de alimentar a los pequeños, daban de comer también al miembro

femenino de la pareja reproductora mientras se hallaba todavía incubando los huevos. Es más, la fusión de individuos no emparentados, ya sea para aumentar el tamaño de la bandada, ya para formar nuevas bandadas, sugiere que las relaciones genéticas no son absolutamente importantes para las formas de cooperación de las abubillas arbóreas.

Se presenta reciprocidad de varios tipos entre compañeros de bandada, ya estén emparentados o no, y ésta parece ser un componente importante del sistema social de los fenicúlidos. Como hemos visto, los miembros no reproductores de la bandada sirven a la vez de asistentes y ayudan en la defensa del territorio. A cambio, estos individuos pueden beneficiarse de dos maneras. Primero, si el asistente sobrevive al reproductor de su propio sexo en la bandada puede heredar la condición de reproductor. Segundo, como resultado de la interacción con los pollos en desarrollo, el asistente gana unos aliados potencialmente valiosos en el caso de que hubiera que emigrar.

Además, cuando un asistente más viejo emigra para sacar partido de un territorio de reproducción vacío, suele ir acompañado de uno o varios socios más jóvenes. Al ayudar a establecer la posición del ave más vieja en la nueva bandada, cabe pensar que las más jóvenes están compensando a la mayor por los servicios que recibieron en el nido. Asimismo, las aves más jóvenes cuidarán de la descendencia de la más vieja, estableciéndose una especie de reciprocidad transgeneracional, puesto que los pollos a los que el ave adulta alimenta pueden, andando el tiempo, ayudar al asistente a acelerar su propio programa de reproducción. La fusión de aves adultas no emparentadas también sugiere reciprocidad; dada la elevada tasa de mortalidad, esta forma de cooperación puede proporcionar con frecuencia la condición reproductora final a todos los implicados.

Finalmente, lo que parece ser la base fundamental que subyace en todo el complejo comportamiento de las abubillas arbóreas es la selección darwiniana individual, es decir, la maximización de las oportunidades reproductoras individuales frente a los demás miembros de la población. Las distintas estrategias de las abubillas arbóreas que hemos descrito se explican de manera más sucinta como un comportamiento que ha evolucionado para maximizar estas oportunidades bajo un conjunto de circunstancias insólitas e impredecibles.

Aneurismas

Las perturbaciones locales del flujo sanguíneo, junto con la tensión hidráulica generada más de 100.000 veces al día por el corazón, son las causas que producen estas dilataciones de la pared arterial

Kaj Johansen

En cada latido, el corazón bombea aproximadamente 70 centímetros cúbicos de sangre a la raíz de la aorta, a una presión de por lo menos 120 milímetros de mercurio, unos 175 gramos por centímetro cuadrado. Después de cada latido, la onda del pulso se propaga a lo largo de las arterias, el sistema ramificado de conductos elásticos y musculares que llevan el oxígeno y los elementos nutritivos a los tejidos. La tensión que se genera por cada latido es aproximadamente de 170.000 dinas por centímetro. A una frecuencia cardíaca media de 70 latidos por minuto esta tensión arterial se repite unas 100.000 veces al día, 2500 millones de veces en el transcurso de una vida.

Por lo general, la integridad estructural de las arterias les permite resistir los efectos destructivos de semejante tensión hidráulica y repetitiva. Sin embargo, en algunas ocasiones la pared arterial cede, y un segmento de la misma se dilata para producir una expansión en forma de globo: un aneurisma. Iniciada la dilatación, tiende a progresar de acuerdo con principios hidráulicos bien conocidos. Si no se trata, el aneurisma puede romperse como consecuencia de la presión interior, causando la muerte o un grave trastorno. Aun cuando no llegue a estallar, el aneurisma puede provocar lesiones al interrumpir el flujo de sangre o por compresión de vasos, órganos o huesos vecinos.

La incidencia de los aneurismas aumenta linealmente con la edad; así, cabe esperar que estas lesiones incidan con mayor frecuencia entre la población de más años. Por otro lado, los avances en las técnicas diagnósticas permiten detectar los aneurismas antes de que produzcan daños irreparables. Se van conociendo mejor las causas que los producen y su historia natural. Se han desarrollado, a la par, nuevas técnicas para tratarlos.

Las arterias son vasos sanguíneos de una gran variedad de diámetros. La aorta es del tamaño de una manguera; a su salida del ventrículo izquierdo del corazón tiene un diámetro interno de aproximadamente 2,5 centímetros y un grosor de pared de unos 2 milímetros. Las arterias más pequeñas, arteriolas que desembocan en los capilares, miden aproximadamente 60 micrometros, 0,06 milímetros, de diámetro. La estructura detallada de las arterias varía de acuerdo con su tamaño, pero en todas ellas la estructura parece apropiada para resistir la tensión del flujo pulsátil repetitivo.

La pared de una arteria consta de tres láminas: la íntima, la media y la adventicia. El componente más interno de la íntima, y que recubre la luz del vaso, es un entretejido continuo de células endoteliales, que ofrecen muy poca resistencia a la presión radial, aunque resisten la tensión longitudinal que supone el flujo sanguíneo. En razón del tamaño de la arteria, el tejido endotelial se halla apoyado sobre mayor o menor cantidad de tejido conectivo. La íntima queda separada de la media por la membrana elástica interna, una suerte de vaina compuesta principalmente por una proteína llamada elastina.

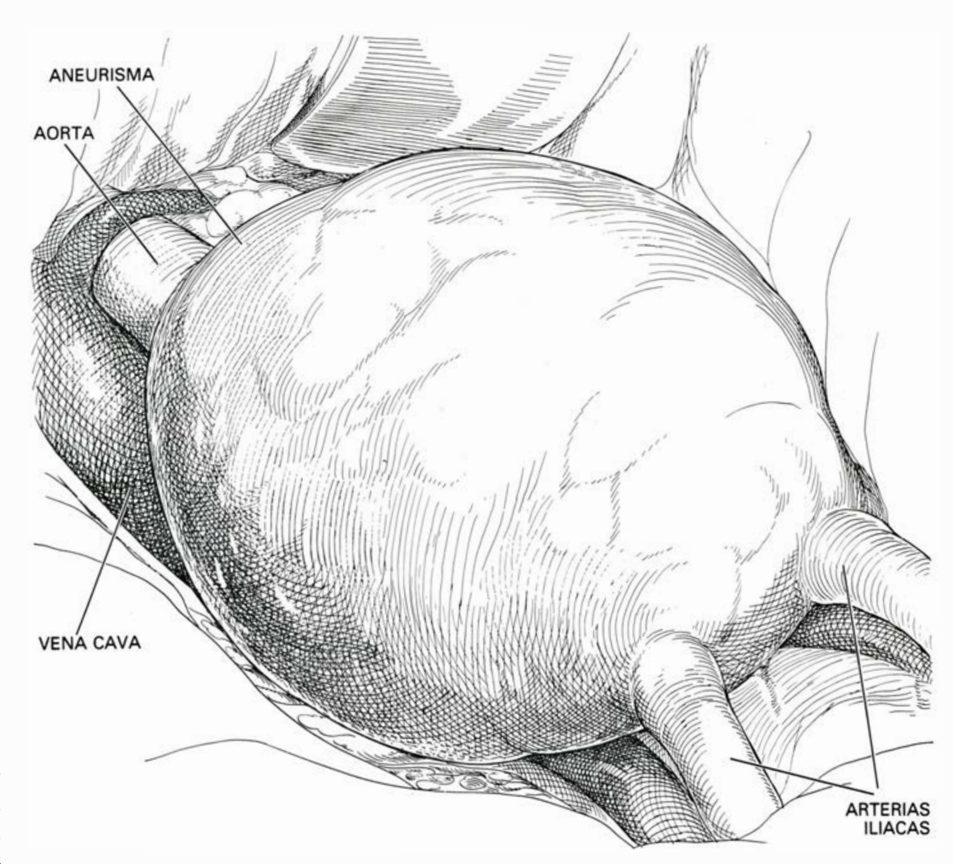
La media, que es responsable de la mayor resistencia y elasticidad de la pared arterial, tiene en la elastina y células musculares lisas sus principales componentes, las proporciones entre las cuales varían según las diferentes arterias. En la aorta y en otras arterias "elásticas" de gran diámetro, la media se halla formada por una membrana elástica dispuesta en varias capas concéntricas, separadas por células musculares lisas y material fibroso. Las paredes elásticas se dilatan para acomodar el volumen de sangre que sale del corazón en cada latido; se contraen después, actuando como una bomba impelente secundaria de la columna de sangre. La capa media de las arterias "musculares" más pequeñas tiene menos componentes elásticos y una mayor proporción de células musculares; estas arterias se contraen y se dilatan en respuesta a señales del sistema simpático y parasimpático y en respuesta a mediadores químicos que circulan en la sangre.

La capa más externa de las arterias es la adventicia, vaina fibrosa compuesta principalmente de colágeno, una proteína del tejido conectivo. Al gozar de débil elasticidad y poseer muy pocas células musculares, no alcanza las propiedades dinámicas de elasticidad de la capa media, pero a ella se debe, en buena parte, la resistencia estática de la pared arterial.

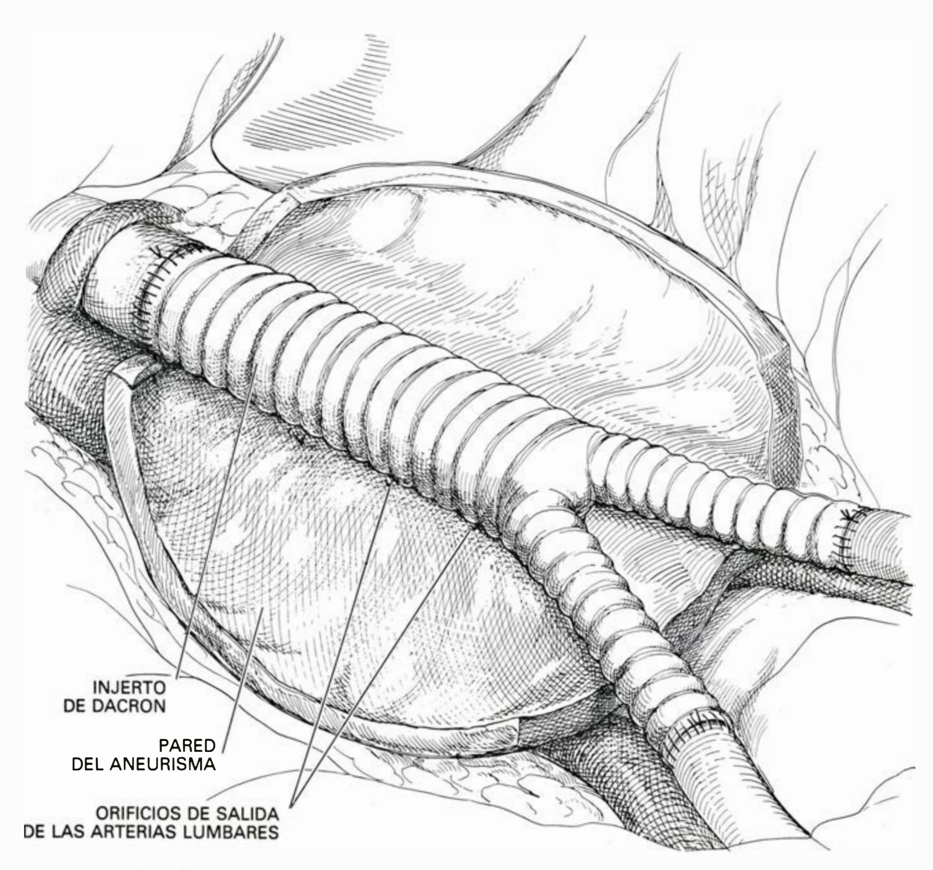
Los verdaderos aneurismas se ven con mayor frecuencia en las arterias mayores, cuya estructura puede darnos indicios sobre la génesis de esas deformaciones. Los elementos elásticos y musculares de la capa media se disponen en forma de láminas elásticas, que envuelven la luz del vaso en una dirección perpendicular a su eje. En esta configuración se hallan en una orientación óptima para resistir la fuerza de distensión producida por la presión pulsátil intraluminal. Las arterias de los individuos que desarrollan aneurismas frecuentemente son ectásicas, es decir, presentan una marcada elongación, hallazgo habitual en los vasos de individuos de edad. La ectasia arterial no constituye de suyo una condición patológica, pero produce, con toda probabilidad, una disrupción de la orientación de las láminas de la capa media. Puesto que el mantenimiento de la estructura vascular depende de un balance entre la presión intraluminal y la capacidad de la pared arterial para resistir la presión, una pérdida de la orientación óptima de los elementos de la media debilita con toda seguridad la pared arterial. Incluso en ausencia de ectasia, con la edad se reduce el contenido elástico de la pared vascular y aumentan las células musculares y los elementos fibrosos. Las arterias viejas son más rígidas que las jóvenes.

Dado que los aneurismas se desarrollan sobre todo en personas de edad avanzada y en arterias ectásicas que ya han sufrido cambios estructurales, debemos aún preguntarnos por qué algunas personas presentan aneurismas y otras no y por qué los aneurismas se ven con mayor frecuencia en puntos específicos de determinadas arterias. Se supone generalmente que los aneurismas son una manifestación de enfermedades arteriales, específicamente de la aterosclerosis. No es difícil comprender las razones que llevan a esta hipótesis. En la aterosclerosis, las células musculares de la capa media proliferan e invaden la íntima. Los lípidos y el tejido cicatricial se depositan sobre la íntima y forman una placa que mengua la luz arterial. Las llamadas placas complicadas se caracterizan por la aparición de calcificación, ulceración y degeneración de la pared vascular. Al observar la deteriorización que se produce en un vaso aterosclerótico es fácil concluir que la aterosclerosis determina una debilitación progresiva de la pared arterial y conduce finalmente a la formación de aneurismas. Además, los pacientes con aneurismas presentan con mucha frecuencia síntomas coexistentes de enfermedad aterosclerótica.

Sin embargo, la conexión entre aterosclerosis y aneurismas puede ser más casual que causal. Los aneurismas acostumbran a formarse en la aorta abdominal, por debajo de las arterias renales y por encima de la bifurcación de la aorta en las dos ramas ilíacas. El mismo segmento de aorta es un sitio común para la aparición de lesiones ateroscleróticas oclusivas. Si un aneurisma fuese simplemente una manifestación tardía de la aterosclerosis, muchos individuos con enfermedad aorto-ilíaca oclusiva desarrollarían aneurismas abdominales aórticos. Sin embargo, no parece ser este el caso. La incidencia de arteriopatía oclusiva en la aorta terminal tiene su máxima incidencia alrededor de los 55 años, mientras que la incidencia de aneurismas aórticos es máxima alrededor de los 70 años; una diferencia demasiado importante como para apoyar la idea de que la enfermedad aterosclerótica causa los aneurismas. Además, la aterosclerosis es una enfermedad sistémica que se manifiesta en muchas arterias. En cambio, los individuos con aneurismas no muestran una incidencia de aterosclerosis generalizada mayor que grupos de pacientes de edades comparables sin aneurismas. Cualquier cirujano vascular puede afirmar que los aneurismas se ven con fre-



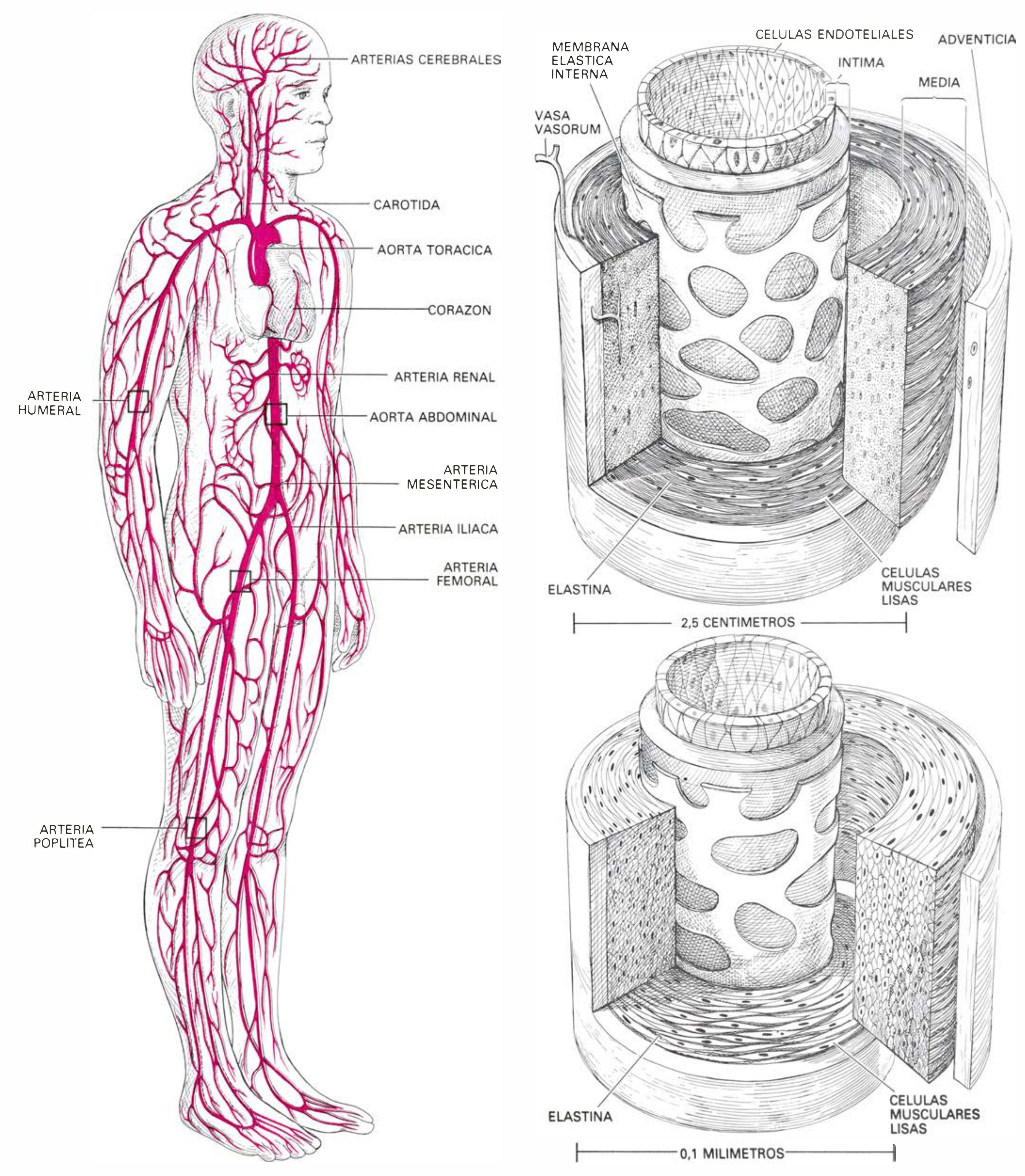
GRAN ANEURISMA de la aorta abdominal dibujado tal como aparece a los ojos del cirujano vascular antes de intervenir. El aneurisma, que es una dilatación de la pared arterial, se halla inmediatamente por encima de la bifurcación de la aorta, que se divide en las dos arterias ilíacas. Se pinzan las arterias por encima y por debajo del aneurisma, se abre el saco dilatado y se sutura el injerto de Dacrón en su interior.



INJERTO PROTESICO, un tubo de Dacrón tejido en forma de *Y*, tal como aparece tras su colocación y sutura a las regiones sanas de la aorta y las arterias ilíacas. Se ha restablecido el flujo sanguíneo.

cuencia en pacientes sin síntomas de lesiones ateroscleróticas. Por último, los lugares donde se localizan con mayor frecuencia los depósitos lipídicos en la aterosclerosis (en las arterias femorales o en la bifurcación de las arterias carótidas del cuello) en muy raras ocasiones son puntos de localización de aneurismas.

Defectos bioquímicos o estructurales hereditarios que no tienen relación alguna con la aterosclerosis quizás influyan en la génesis de los aneurismas. David W. Rowe, Ermona B. McGoodwin y sus colegas, en el Argonne National Laboratory, han demostrado que los aneurismas aórticos se desarrollan espontáneamente en una especie de ratón con una mutación específica del cromosoma *X* que determina la aparición de degeneración arterial en la gran mayoría de los animales portadores. La mutación parece inhibir la formación de un puente molecular entre la elastina y el colágeno mediante la lisil oxidasa, un enzima que contiene cobre. La insuficiente formación de puentes es-



LOCALIZACIONES MAS FRECUENTES de los aneurismas (*cuadrados negros*) en un diagrama del árbol arterial. Los sitios más comunes se hallan inmediatamente por encima de las bifurcaciones de las principales arterias.

ESTRUCTURA DE LA PARED ARTERIAL, que varía con el tamaño de la arteria y con su función. Los esquemas muestran (a diferentes escalas) la pared de la aorta (*arriba*) y la de una pequeña arteria muscular (*abajo*).

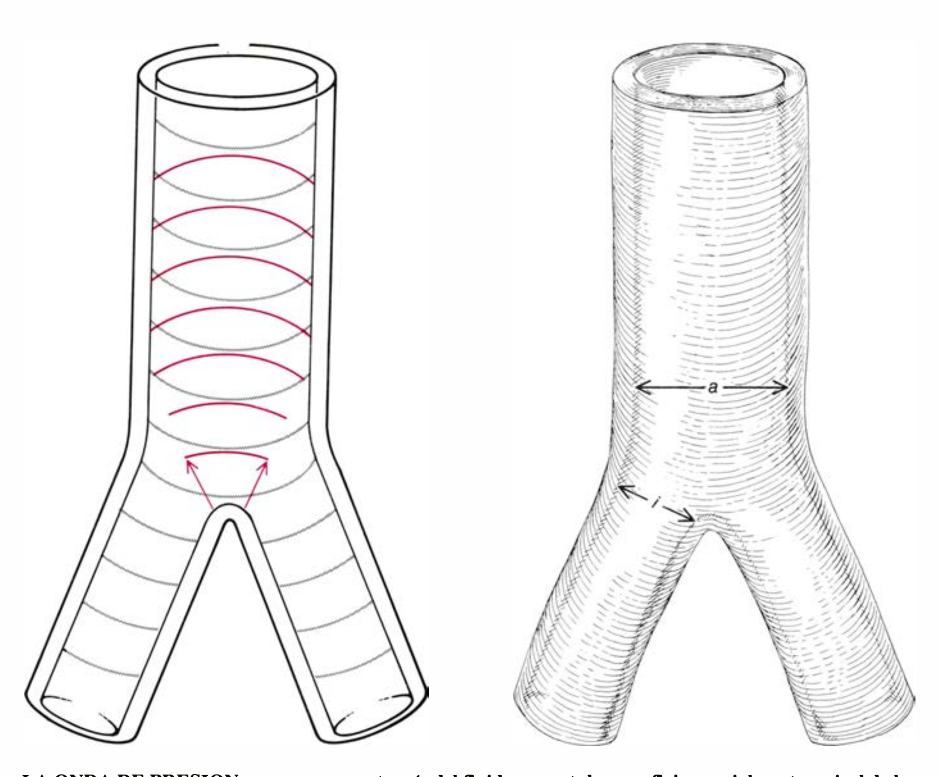
tructurales entre la elastina y el colágeno puede determinar la debilitación de la pared arterial. En los individuos humanos se ha observado una rara enfermedad, conocida por el síndrome de Menkes, que se caracteriza por la presencia de anomalías en la estructura del colágeno y por malabsorción del cobre, y que también parece heredarse a través del cromosoma X; a veces se observan aneurismas arteriales en niños afectos de esa enfermedad.

Los estudios realizados por Ronald W. Busuttil, de la Facultad de Medicina de la Universidad de California en los Angeles, y por Philip B. Dobrin, de la Facultad de Medicina Stritch de la Universidad Loyola, sugieren que en el debilitamiento de la pared arterial podría intervenir la disolución enzimática de las subunidades estructurales de esa pared. En pacientes de aneurisma, los enzimas que degeneran el colágeno y la elastina quizá se hallen en cantidades mayores o en proporciones anormales.

M. David Tilson, de la Facultad de Medicina de la Universidad de Yale, ha probado la existencia de correlación entre la arteriomegalia, es decir, presencia congénita de arterias de un tamaño mayor de lo normal, y aneurismas. Nuestro grupo, del Hospital Clínico Harborview de la Facultad de Medicina de la Universidad de Washington, se encuentra empeñado en un estudio de las familias de pacientes que presentan aneurismas aórticos. Por ahora, nuestra impresión es que la incidencia de esta enfermedad en parientes próximos de los enfermos es mayor de lo que podría esperarse por azar, lo que apunta hacia la existencia de un rasgo hereditario (que se expresaría de forma tardía en la vida) causante del deterioro y posterior dilatación del vaso.

Pudiera ocurrir que ni la enfermedad arterial ni los defectos congénitos fueran la causa primaria que originase los aneurismas. Las lesiones podrían generarse simplemente en sitios muy vulnerables tras muchos años de stress hidráulico pulsátil. El aneurisma más común, como hemos mencionado, se ubica cerca de la porción más distal de la aorta abdominal, inmediatamente por encima de su bifurcación en las arterias ilíacas. Trabajos experimentales en varios centros sugieren que la incidencia del deterioro de la pared arterial en ese lugar puede explicarse por las perturbaciones locales específicas del flujo sanguíneo.

Una onda que se transmita a través de un tubo se reflejará parcialmente en las bifurcaciones. Ray G. Gosling y sus colegas, del Hospital Infantil de Londres, han demostrado en perros y otros animales experimentales, así como en estudios necrópsicos de arterias humanas, que la naturaleza y la magnitud de la onda que se refleja a partir de la bifurcación aórtica pueden hallarse relacionadas matemáticamente con la razón que media entre el área transeccional de las dos arterias ilíacas y el área transeccional de la aorta. La razón entre dichas áreas se hace progresivamente más pequeña con la edad, a medida que la aorta comienza a volverse más fláccida; conforme disminuye dicha razón, aumenta la intensidad de la onda reflejada desde la bifurcación aórtica. Cualquier dilatación inicial de la aorta (debida a arteriomegalia, ectasia o a una reducción de la elasticidad arterial relacionada con la edad) disminuye todavía más la proporción entre ambas áreas, incrementándose la reflectancia de las ondas y creciendo la tensión hidráulica de la pared aórtica. En 1954, Arthur B. Voorhees, Jr., y Arthur H. Blakemore, de la Facultad de Medicina y Cirugía de la Universidad de Columbia, sugirieron que un efecto en "martillo de agua", que se genera cuando la onda reflejada choca con la presión del pulso del latido siguiente, determina la aparición de una tensión adicional sobre la pared aórtica.



LA ONDA DE PRESION que se mueve a través del fluido en un tubo se refleja parcialmente a nivel de las bifurcaciones (*izquierda*), generando un stress lateral en la pared, inmediatamente por encima de la bifurcación. Tales ondas reflejadas pueden ser parcialmente responsables de la formación de aneurismas justo antes de la bifurcación de la aorta (*derecha*). La magnitud de la onda reflejada se relaciona de forma inversa con la razón entre el área de la luz de las dos arterias ilíacas y el área seccional de la aorta, es decir 2*i*/*a*. La razón de las áreas tiende a disminuir con la edad, a medida que la aorta se hace más rígida, lo que produce un aumento de la intensidad de las ondas reflejadas y favorece la aparición de aneurismas.

Otro principio hidráulico podría ayudar a la progresión de los aneurismas. Cuando el calibre de un tubo aumenta gradualmente (como sucede con la luz aórtica en un aneurisma incipiente), el fluido cercano a la pared pierde velocidad, generándose un flujo turbulento. La turbulencia puede inducir la aparición de vibraciones en la pared. Margot R. Roach, de la Facultad de Medicina de la Universidad de Western Ontario, ha demostrado que la vibración de baja frecuencia de la pared arterial (hasta 200 ciclos por segundo en las arterias ilíacas humanas) puede dañar severamente los componentes estructurales de la pared aórtica. Desde el punto de vista anatómico, la importancia de estos diversos mecanismos hidráulicos surge de la ubicación habitual de los aneurismas: inmediatamente por encima de la bifurcación de las arterias mayores. Puntos susceptibles para la generación de estas tensiones son, además de la aorta distal, las arterias ilíacas y femorales de la pelvis y de la porción superior del muslo y la arteria poplíteal, detrás de la rodilla.

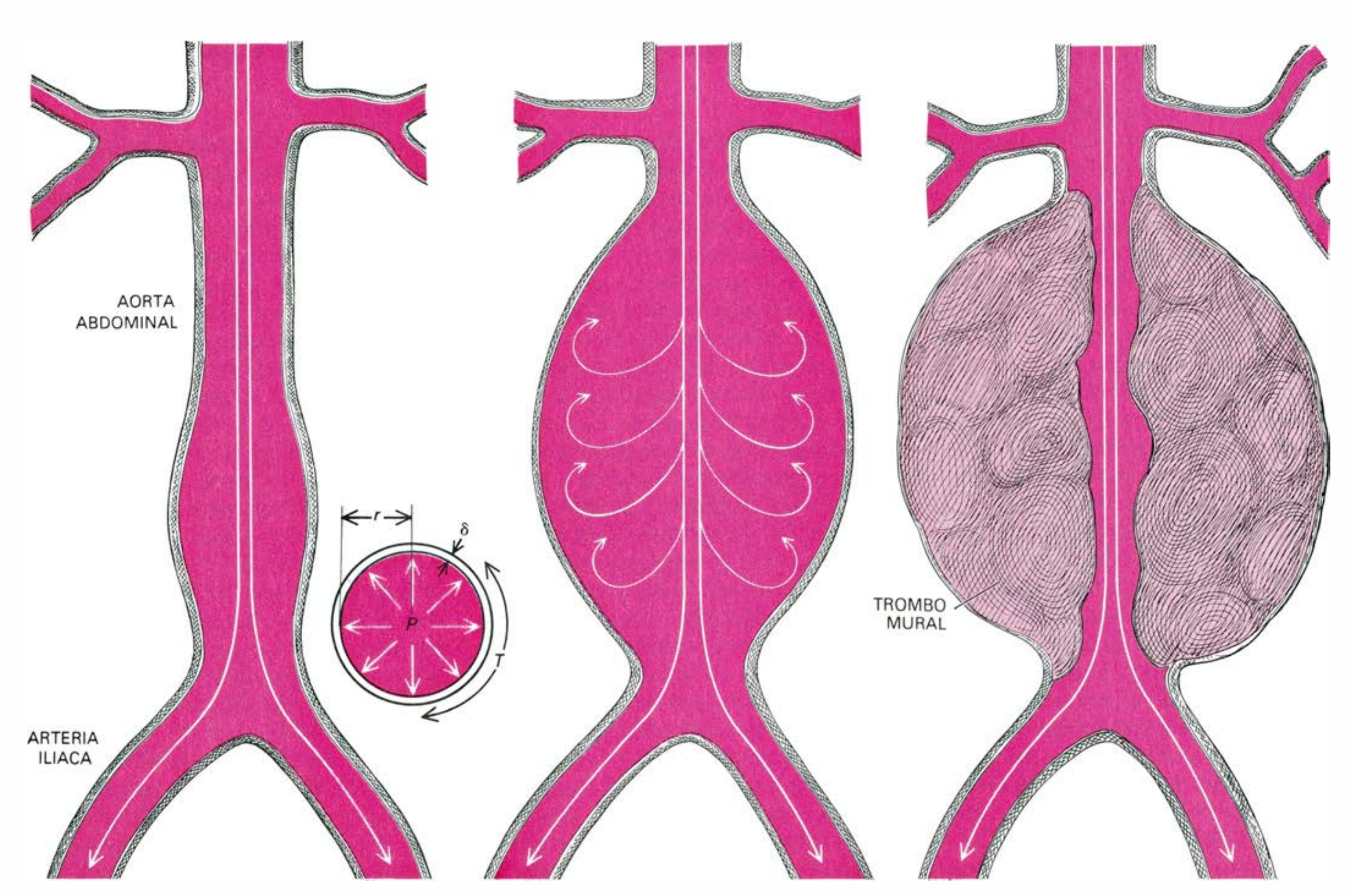
Si nos atenemos a las pruebas que se van acumulando, cada vez se ve más claro que la sensibilidad de la aorta abdominal al deterioro estructural obedece a un inadecuado aporte de oxígeno y de elementos nutritivos a esta zona de

la arteria y a alguna deficiencia estructural. El segmento terminal de la aorta tiene menos vasa vasorum (arteriolas que llevan los nutrientes a las porciones más externas de los vasos arteriales) que otras arterias. Las investigaciones de Seymour Glagov y sus colegas, de la Facultad de Medicina Pritzker de la Universidad de Chicago, indican que el segmento de la aorta abdominal en los humanos puede padecer una constitución deficiente de sus envolturas músculo-elásticas. Entre las arterias renales y la bifurcación de la aorta se identifican 28 láminas elásticas en la capa media, mientras que los estudios comparativos con animales sugieren que debería esperarse encontrar por lo menos 35. En virtud de esta menor cantidad de láminas músculo-elásticas, la tensión que debe soportar cada una de ellas es más del doble en la aorta abdominal (donde los aneurismas son especialmente comunes) que en la aorta torácica del hombre (donde los aneurismas son muy raros).

Algunas enfermedades vasculares específicas se asocian con la formación de aneurismas. La displasia fibromuscular, una enfermedad que parece afectar a la media, confiere a la arteria un aspecto arrosariado (con áreas estrechadas entre dilataciones aneurismáticas). Tales lesiones menudean en las arterias renales y ocasionalmente en las carótidas; llegan a causar hipertensión renovascular o accidentes cerebrovasculares.

Surgen pequeños aneurismas de las arterias cerebrales en localizaciones (generalmente puntos de bifurcación) donde la media se halla ausente o es deficitaria. Estos aneurismas pueden estallar, determinando la aparición de hemorragias en o alrededor del cerebro, lo que provoca lesiones neurológicas. La enfermedad de Kawasaki (mal que afecta a recién nacidos y que puede tener un origen vírico) se caracteriza por la aparición de aneurismas en las arterias coronarias. Los obreros que trabajan con máquinas que vibran, martillos neumáticos o sierras de cadena por ejemplo, desarrollan aneurismas de las pequeñas arterias de la mano. De todas formas, en la gran mayoría de casos la cadena de influencias que determinan un aneurisma parece comenzar con la presencia de un defecto estructural local, asociado quizá con un aporte inadecuado de sangre a la pared arterial. Las anomalías hidráulicas que se producen al generarse el flujo turbulento o el reflejo de las ondas del pulso aumentan de una forma extraordinaria la tensión parietal. El resultado es una dilatación que, una vez iniciada, progresa inexorable y determina la formación de un aneurisma.

Se caracteriza el aneurisma propiamente dicho por la dilatación de las tres capas de la pared arterial. Un pseudoaneurisma, que es una masa pulsátil que presenta muchos de los hallazgos clínicos peculiares de los aneurismas, puede producirse cuando la sangre rompe, bajo presión, una o dos de las capas arteriales, o bien cuando se produce una rotura completa de la pared arterial y se forma un hematoma. La causa más frecuente de un pseudoaneurisma son las lesiones penetrantes: heridas por arma blanca o por bala, daño provocado por una aguja o un catéter introducido



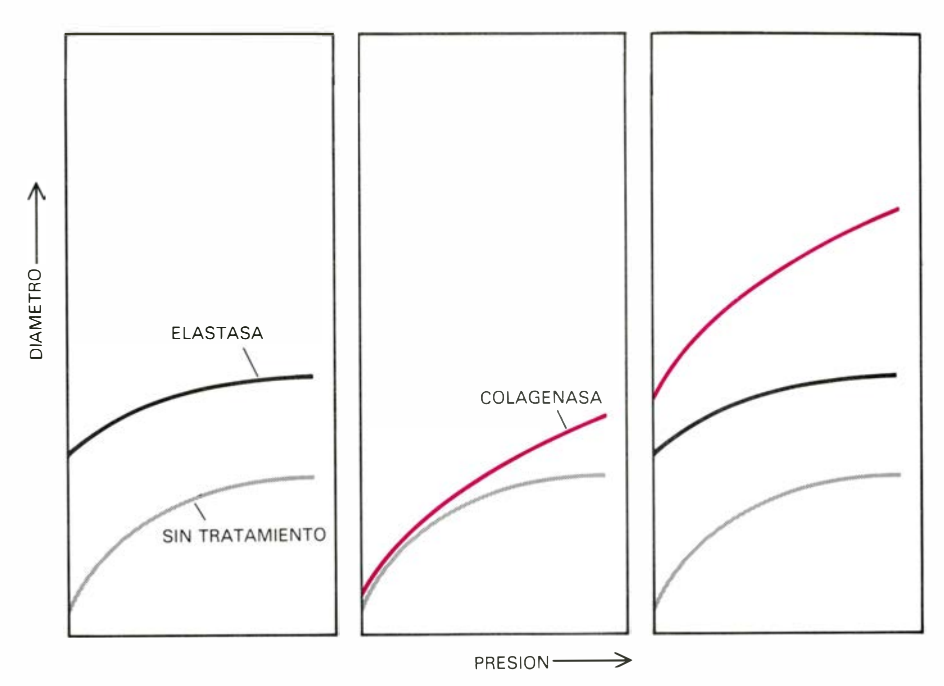
ANEURISMA DE LA AORTA ABDOMINAL. Una vez iniciado, tiende a aumentar inexorablemente de tamaño. La tensión (*T*) de la pared varía directamente con el producto de la presión intraluminal (*P*) y el radio (*r*) de la luz, e inversamente con el grosor (δ) de la pared: *T* es proporcional a *Pr*/δ. Cualquier dilatación, por pequeña que sea (*izquierda*), aumenta el radio y disminuye el grosor de la pared, lo que genera un aumento de la tensión y el agrandamiento del aneurisma. A medida que la arteria se dilata, el flujo sanguíneo próximo a la pared se torna más lento, lo que crea un régimen turbulento (*centro*); el flujo turbulento determina vibraciones de la pared (hasta 200 ciclos por segundo en las arterias ilíacas humanas) que la debilitan aún más. Ambos factores, la turbulencia y las irregularidades en el interior de la pared lesionada, promueven la formación de trombos murales: la agregación de plaquetas, células sanguíneas y detritus celulares, entretejidos con la proteína llamada fibrina. El trombo puede reducir de forma muy significativa el diámetro de la luz (*derecha*). También puede fragmentarse, formar émbolos y producir oclusiones de arterias menores por debajo del nivel del aneurisma.

en una arteria en un procedimiento diagnóstico terapéutico, o una lesión producida por infección o por un defecto en la sutura de una arteria en una operación quirúrgica vascular.

Desarrollado el aneurisma, crece la tensión de su pared de acuerdo con el principio hidráulico fundamental enunciado hace más de 200 años por el astrónomo y matemático francés Pierre Simon de Laplace. El principio relaciona la tensión de la pared con la presión y el tamaño del vaso. Explica, entre otras cosas, el estallido de un globo o de una pompa de jabón. La ley de Laplace afirma que la tensión (T) en la pared de un recipiente esférico es proporcional al producto de la presión intraluminal (P) y del radio (r) de la esfera: T es proporcional a Pr. Dicha ley se aplica a la esfera de pared infinitamente delgada. D. Eugene Strandness, Jr., y David S. Sumner, de la Facultad de Medicina de la Universidad de Washington, han demostrado que el grosor de la pared (δ) debe desempeñar algún papel en la distensión y rotura final de un aneurisma. (Un globo de paredes gruesas resiste mejor la presión que otro de paredes delgadas.) La tensión en la pared arterial, en consecuencia, varía directamente con el producto de la presión y el radio y de forma inversa con el grosor de dicha pared: T es proporcional a Pr/δ.

Dado que el volumen de la pared arterial presumiblemente permanece constante, el grosor de la pared debe disminuir a medida que se agranda el aneurisma. El valor de $(Pr)/\delta$, en consecuencia, aumenta deprisa: la tensión de la pared en un aneurisma de 6 centímetros de diámetro es 12 veces mayor que la tensión en la pared de la aorta abdominal normal con un diámetro interno de 2 centímetros. El aumento de la tensión determina una disminución aún mayor del grosor de la pared y un aumento del tamaño del aneurisma, que, a su vez, incrementa la tensión. Este círculo vicioso, exacerbado quizá por factores clínicos, como la hipertensión (un aumento de P), termina por romper un aneurisma.

A medida que se produce un aneurisma, su pared se va recubriendo de un trombo mural: es decir, un agregado de células sanguíneas, plaquetas, proteínas de la sangre y detritus celulares. El trombo crece a medida que se deposita en él material procedente de la sangre; con el tiempo, el canal que permite el paso de la sangre a través de la zona del aneurisma se torna del tamaño de la luz que existía antes del desarrollo de éste. Se ha sugerido que la formación de



EFECTO DE DOS ENZIMAS en las arterias carótidas de los perros; lo estudió Philip B. Dobrin, de la Facultad de Medicina Stritch de la Universidad Loyola. Para ello trató las arterias con elastasa y con colagenasa, enzimas que degradan la elastina y el colágeno, respectivamente, y midió el diámetro de los vasos a medida que se aumentaba la presión intraluminal. La elastasa dilató la arteria (*izquierda*) pero no la rompió. La colagenasa produjo sólo una leve dilatación (*centro*) pero provocó la rotura arterial. El tratamiento con elastasa seguida de colagenasa determinó dilatación arterial y ruptura (*derecha*).

trombos es un intento por parte del organismo de aumentar el grosor de la pared y relajar así la tensión de la pared y la probabilidad de rotura del aneurisma. El trombo es en realidad una masa gelatinosa. Y puesto que la presión se halla equilibrada en todos los puntos de un sistema hidráulico, la tensión de la pared es la misma exista o no un trombo intraluminal. En consecuencia, la ayuda que ofrece la presencia de un trombo mural para evitar la rotura de un aneurisma es sólo marginal.

Los aneurismas no son raros. En un estudio necrópsico practicado en Suecia se encontraron aneurismas de la aorta abdominal en casi un 2 por ciento de las autopsias. Aneurismas de todo tipo, en la aorta y en otras arterias, se aprecian en hasta un 10 por ciento de todos los exámenes necrópsicos.

En exámenes físicos rutinarios, se encuentra uno a veces con aneurismas de la aorta abdominal al palpar en el abdomen una masa pulsátil. En otros casos, se diagnostica su presencia a partir de las complicaciones propias de los aneurismas. Pensemos, por ejemplo, en la embolización periférica de pequeñas partes del trombo mural. Estos émbolos pueden localizarse en arterias más pequeñas e interrumpir de forma súbita el flujo sanguíneo. El trombo mural, al crecer, estenosa la luz de la arteria hasta interrumpir el flujo sanguíneo. La trombosis de un aneurisma es una complicación particularmente seria a nivel de las arterias poplíteas, por detrás de la rodilla. A este nivel pocas arterias colaterales pueden suplir el flujo distal más allá de la oclusión arterial; en consecuencia, la pérdida de la extremidad es resultado frecuente de la trombosis de un aneurisma poplíteo. En la aorta, o a nivel de las arterias ilíacas, un aneurisma puede agrandarse hasta el punto de comprimir vísceras huecas circundantes, tales como el uréter; puede erosionar la columna vertebral o un segmento del intestino.

La manifestación más grave de la enfermedad aneurismática es la rotura, más frecuente en los aneurismas de gran tamaño de la aorta abdominal. Varios estudios, entre ellos uno realizado por Eugene F. Bernstein y George R. Leopold, de la Facultad de Medicina de la Universidad de California en San Diego, han mostrado que la gran mayoría de aneurismas aumentan de tamaño lentamente, con una dilatación anual de 0,4 centímetros. Sin embargo, los aneurismas continúan creciendo y, como predice la ley de Laplace, cuanto mayor sea el tamaño del aneurisma más probabilidades existen de que estalle. Lesiones de sólo 4 centímetros de diámetro pueden romperse. Una incidencia significativa de roturas aneurismáticas puede esperarse una vez el diámetro supere los cinco centímetros. D. Emerick Szilagyi ha mostrado que la probabilidad de rotura para las lesiones de más de 6 centímetros de diámetro rebasa el 50 por ciento.

Si el aneurisma se rompe hacia el peritoneo (la membrana que envuelve las vísceras abdominales) el paciente puede perder rápidamente grandes cantidades de sangre en la cavidad abdominal; el shock irreversible, paro cardíaco y la muerte se producen a los pocos minutos de la rotura del aneurisma. Lo normal es que los aneurismas se rompan en el retroperitoneo, detrás de la cavidad abdominal, donde la resistencia y la presión debida a los músculos y otros tejidos blandos pueden, de forma temporal, cohibir la hemorragia que proviene de la arteria perforada. En esta situación, el paciente busca asistencia médica debido al fuerte dolor en la región de la espalda, o se le conduce a la sala de urgencias del hospital por presentar un estado de shock.

La pérdida de sangre debe reemplazarse inmediatamente mientras se inician los preparativos para una operación de urgencia, en que la aorta se pinzará justo por encima del aneurisma. Si se efectúa la intervención antes de que el enfermo haya perdido demasiada sangre y de que los órganos vitales hayan sufrido un daño irreversible, sobrevivirá. La mayoría de pacientes con rotura de un aneurisma son de edad avanzada, razón por la que suelen presentar otros problemas médicos asociados, especialmente cardiopatías. Si el individuo sobrevive en estas circunstancias a la situación de emergencia, quizá muera de un infarto de miocardio, de un accidente cerebro-vascular o por insuficiencia renal. De un estudio realizado por nuestro grupo en el Hospital Clínico Harborview se desprende que un tercio de los pacientes que sobreviven a una rotura de aneurisma pueden presentar diversos grados de gangrena del intestino grueso, resultado quizá de la pérdida masiva de sangre. La mortalidad por rotura de un aneurisma aórtico se cifra en torno al 50 por ciento, y algunas series indican que puede llegar al 90 por ciento de los pacientes que sufrieron shock.

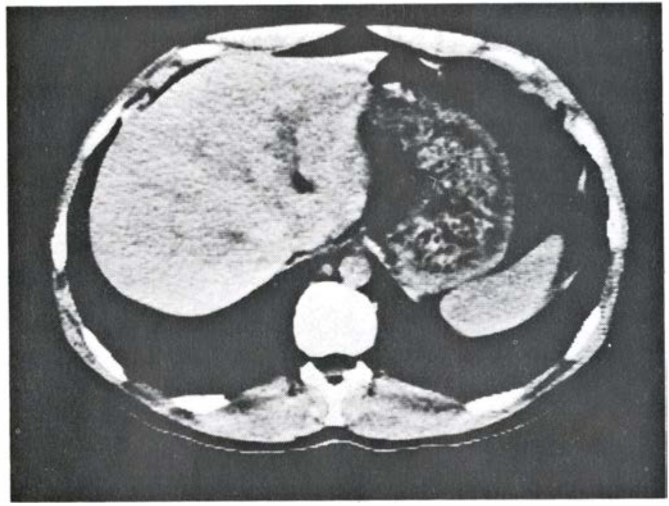
Desde finales de la década de 1940 se sabe cómo identificar la presencia de una aneurisma y estimar su tamaño mediante el recurso a la arteriografía. Se trata de una técnica por la que se visualizan las arterias en una radiografía después de inyectar un contraste opaco a los rayos X en la arteria a través de una aguja o de un catéter. La técnica requiere punción arterial; la presencia de trombos murales en el interior del aneurisma puede producir, por ende, errores significativos a la hora de estimar el tamaño real de éste.

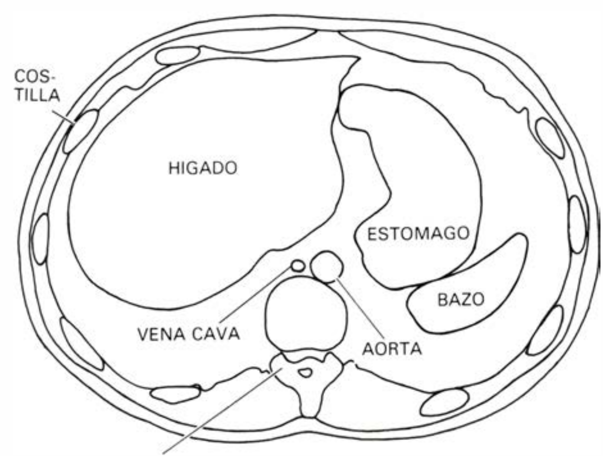
La práctica médica empieza a servirse de otras dos técnicas no invasivas de gran importancia: los ultrasonidos y la tomografía axial computarizada (TAC). La imagen ultrasonográfica se forma por la reflectividad diferencial de los distintos tejidos a los sonidos de alta frecuencia. Una exploración TAC es una sección transversal de un plano del cuerpo que recoge los detalles de las estructuras blandas y del hueso y que se construye mediante el tratamiento por ordenador de las imágenes de rayos X. Las imágenes se obtienen haciendo girar una fuente de rayos X alrededor del plano a investigar. Estos métodos modernos permiten medir el tamaño de un aneurisma pequeño y asintomático con precisa exactitud; cabe también realizar determinaciones seriadas y, en consecuencia, evaluar el desarrollo de la enfermedad.

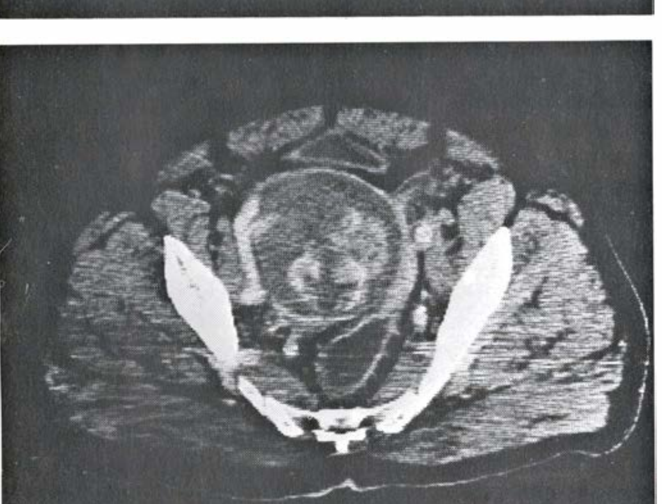
La mortalidad asociada al tratamiento quirúrgico electivo de un aneurisma aórtico ha disminuido hasta el uno o el dos por ciento en algunas series, aun cuando se incluyen en ellas pacientes octogenarios. La incidencia de complicaciones serias es, asimismo, baja. Los individuos que han superado una intervención tienen aproximadamente la misma esperanza de vida que un grupo coetáneo suyo que nunca padeciera aneurismas. Estos resultados, sumados a la elevada mortalidad observada en los casos de rotura aneurismática y al hecho de que no hay tratamiento efectivo fuera del quirúrgico para los aneurismas, apoyan la intervención electiva de esta enfermedad. El tratamiento quirúrgico electivo debe realizarse en cualquier aneurisma que supere los cinco centímetros de diámetro, que sea doloroso a la palpación o bien que dé síntomas tales como dolor de espalda. Hasta los pequeños aneurismas asintomáticos deberían ser objeto de intervención, a menos que el paciente sufra enfermedades coexistentes que aumenten el riesgo quirúrgico.

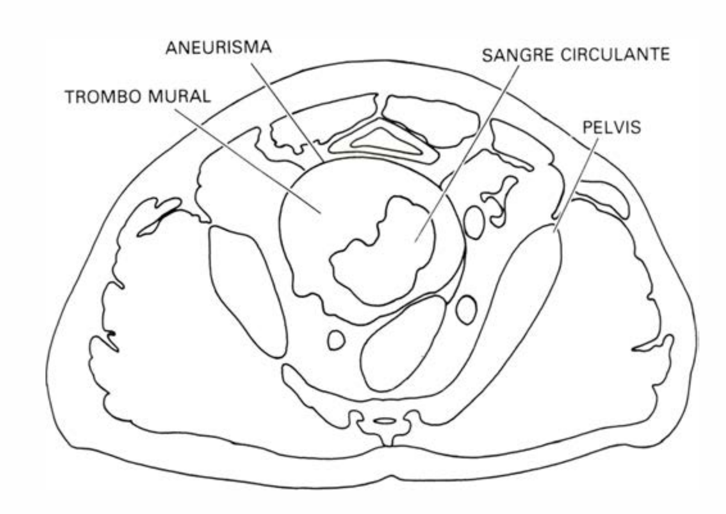
La naturaleza letal de los aneurismas se conoce desde hace muchos años. Hasta las últimas décadas la causa más común de su formación probablemente no era la tensión o el stress mecánico de las arterias al envejecer, sino la infección (debida a la sífilis o a la colonización bacteriana de válvulas cardíacas infectadas) o debido a lesiones arteriales, a menudo consecuencia de la extendida práctica del sangrado. El tratamiento quirúrgico de los aneurismas de la arteria poplítea y de las arterias braquiales se intentó ya en el siglo XVIII. La arteria se ligaba por encima y por debajo del aneurisma, lo que casi inevitablemente comportaba la pérdida de la extremidad.

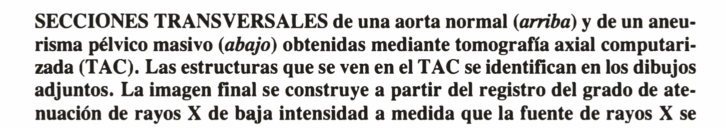
ARTERIOGRAFIAS que delinean el contorno de dos aneurismas aórticos abdominales, uno de los cuales (*derecha*) es una lesión que aparece en una aorta ectásica y elongada. Las radiografías se tomaron después de la introducción en el interior de la aorta de un medio de contraste (una sustancia opaca a los rayos X). En ambas fotografías, el catéter a través del que se inyectó el medio de contraste aparece como una línea blanca de arriba abajo. El catéter se introdujo a través de la arteria femoral y se avanzó por la arteria ilíaca hasta la aorta, por encima del aneurisma, donde se efectuó la inyección del medio de contraste.



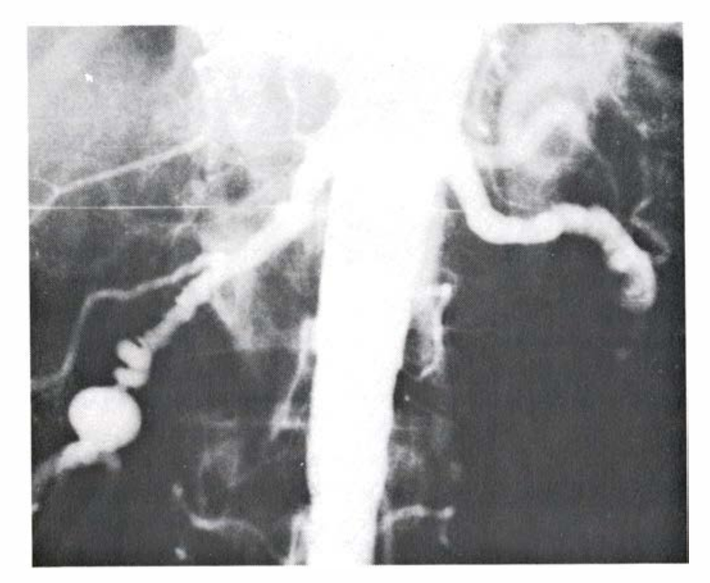






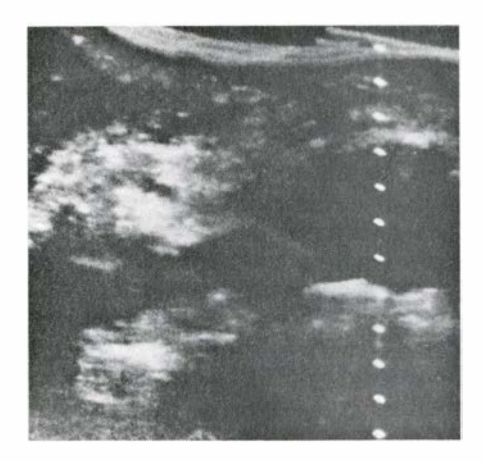


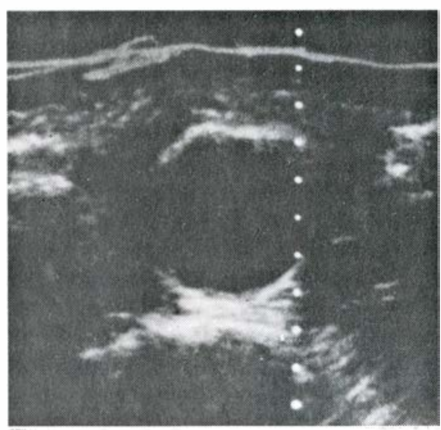
SECCIONES TRANSVERSALES de una aorta normal (*arriba*) y de un aneurisma pélvico masivo (*abajo*) obtenidas mediante tomografía axial computarizada (TAC). Las estructuras que se ven en el TAC se identifican en los dibujos adjuntos. La imagen final se construye a partir del registro del grado de atenuación de rayos X de baja intensidad a medida que la fuente de rayos X se rota alrededor del individuo. Las dos imágenes se tomaron a diferentes niveles: la aorta normal se registró en la porción más alta del abdomen, mientras que la sección que revela el aneurisma se tomó a nivel de la pelvis. Las tomografías representan secciones transversales del cuerpo vistas por debajo. Esta moderna técnica permite distinguir un trombo de la sangre en movimiento.





LOS ANEURISMAS pueden producirse como consecuencia de lesiones vasculares específicas. La displasia fibromuscular (*izquierda*) es una afectación de la media (la capa que confiere mayor consistencia a la pared arterial). Esta enfermedad produce un rosario de pequeños aneurismas separado por lesiones estenóticas. La arteriografía muestra tales aneurismas en las arterias renales. Un aneurisma de una arteria cerebral (*derecha*) puede producirse cuando la media de una de esas arterias es muy delgada o no existe. Se muestra un gran aneurisma en la arteria carótida interna, en la base del cerebro.





ECOGRAFIAS abdominales, que permiten estimar el tamaño y la morfología de un aneurisma con toda exactitud sin acudir a la inyección de medio de contraste. Se proyectan sobre el cuerpo ondas de sonido de muy alta frecuencia y se registra la intensidad de los ecos que inducen; los tejidos se distinguen unos de otros por sus diferentes capacidades para reflejar las ondas sónicas. Las figuras representan la imagen longitudinal (*izquierda*) y transversal (*derecha*) de un aneurisma de la aorta abdominal, y se obtuvieron desplazando la fuente de sonido a lo largo de la línea media y de lado a lado del abdomen. Los puntos blancos son la escala y se hallan separados por un centímetro. (La ilustración es de Kaj Johansen.)

En el siglo XIX se intentaba inhibir el aumento de tamaño de un aneurisma mediante maniobras que favorecían la formación de trombos intraluminales, bien insertando agujas en su interior o bien aplicando vendajes compresivos alrededor de la pierna o del abdomen. En 1864, un cirujano inglés introdujo más de 20 metros de alambre de hierro muy fino en un aneurisma para inducir la trombosis. La electrocoagulación de los aneurismas la intentó por primera vez, en 1879, el médico italiano Corradi, que introdujo 42 centímetros de alambre conductor en el aneurisma e hizo circular a su través la corriente eléctrica. La técnica se aplicó de forma bastante generalizada durante muchos años y con éxitos esporádicos.

La reparación real de un aneurisma llegó cuando los adelantos técnicos permitieron a los cirujanos reemplazar o introducir injertos alrededor de los vasos sanguíneos afectados. Los nuevos desarrollos contemplaban la aplicación clínica de heparina para evitar la coagulación de los vasos pinzados, la disponibilidad de la arteriografía y una mejor comprensión de los principios básicos de la cirugía vascular. La primera intervención de reemplazo de la aorta distal (que se realizó para tratar una aterosclerosis oclusiva y no un aneurisma) la practicó Jacques Oudot, en el Hospital Broussais de París, en 1951. Ese mismo año, y en el mismo centro, Charles Dubost, Michel Allary y Nicolas Oeconomos extirparon completamente un aneurisma de la aorta abdominal y lo sustituyeron por una arteria extraída de un cadáver. Michael E. DeBakey y Denton A. Cooley, a la sazón en el Baylor University College of Medicine, repararon un aneurisma de la aorta torácica en 1953. Un año más tarde, Charles G. Rob, del Hospital St. Mary de Londres, utilizó el primer injerto sintético para resecar un aneurisma de la aorta abdominal. Para ello se sirvió de un tubo de tejido obtenido de una camisa cosido con una máquina de coser familiar. Pronto se dispuso de injertos protésicos, generalmente hechos de Dacrón; a mediados de la década de los 60, se empleaban de forma rutinaria.

Sin embargo, la mortalidad operatoria era inaceptablemente alta. El tejido que forma la pared de un aneurisma aórtico suele fusionarse a las estructuras vecinas. En muchos casos era difícil resecar el aneurisma sin dañar tales estructuras, en particular la vena cava inferior, la gran vena que recoge la sangre de las extremidades inferiores y del tronco y la devuelve al corazón. Muchos cirujanos vasculares adoptaron entonces una técnica de reparación intravascular, merced a la cual la arteria se pinzaba temporalmente por encima y por debajo del aneurisma y se abría el saco aneurismático dilatado. Se suturaba entonces el injerto protésico a los segmentos sanos y no dilatados de la arteria. Este procedimiento evitaba la necesidad de una disección extensa fuera del aneurisma y obviaba la lesión de las estructuras vecinas.

Los aneurismas de las arterias pequeñas del abdomen y de las arterias de la cabeza y de las extremidades presentan menos riesgos de rotura que los de la aorta, pero son mayores los riesgos de trombosis o de embolización, por lo que también deben tratarse quirúrgicamente una vez diagnosticados. Si la localización del aneurisma tiene suficientes vasos colaterales paralelos al enfermo, basta con ligar la dilatación. En otros casos habrá que crear un puente protésico; generalmente el injerto es un segmento de la vena safena extraída de la pierna del enfermo o bien una pieza de tubo de material plástico. Los aneurismas originados a raíz de procesos infecciosos plantean especiales dificultades para su tratamiento quirúrgico. Frecuentemente resultan de heridas penetrantes; verbigracia: cuando la arteria se punciona para inyección de drogas o en el curso de procedimientos diagnósticos tales como la arteriografía. El área de la infección debe desbridarse ampliamente y drenarse; a continuación se ha de restablecer el flujo arterial más allá del aneurisma mediante la creación de un puente a través de tejido sano.

El tratamiento de los aneurismas de las arterias cerebrales puede encerrar cierto grado de complejidad, habida cuenta de la dificultad de llegar hasta ellos y de que la gran mayoría de pacientes en que se plantea han sufrido con anterioridad lesiones neurológicas graves, resultado de una súbita y catastrófica hemorragia intracraneal. Un tratamiento quirúrgico directo –el más común de los utilizados es un intento de ligar la base del aneurisma– constituye el tratamiento apropiado en la mayoría de los casos. También puede recurrirse a técnicas quirúrgicas alternativas. Así, disminuir la presión intraluminal a nivel de la arteria que tiene el aneurisma mediante una oclusión gradual de la arteria carótida en el cuello. Un pequeño globo, introducido mediante un catéter y conducido hasta el interior del aneurisma puede inducir trombosis; en algunos casos puede inyectarse un adhesivo especial en el interior de la cavidad del aneurisma. John F. Alksne, de la Facultad de Medicina de la Universidad de California en San Diego, ha desarrollado una técnica por la que el aneurisma se perfora con una aguja que se introduce a través de una trepanación; la aguja introduce luego fragmentos muy finos de hierro mientras se aplica un campo magnético. Este procedimiento tiende a inducir trombosis en el aneurisma.

Un mejor conocimiento de los factores que predisponen a los individuos al desarrollo de los aneurismas y el perfeccionamiento de las técnicas diagnósticas potenciarán la capacidad de los cirujanos para tratar y diagnósticar estas peligrosas lesiones.

Temas metamágicos

La fecunda progenie del Cubo: esferas, pirámides, dodecaedros y sabe Dios qué

Douglas R. Hofstadter

¡Oh, suerte maldita!
¡Que haya nacido yo para ponerlo en orden!
–*Hamlet*, Acto I, Escena V

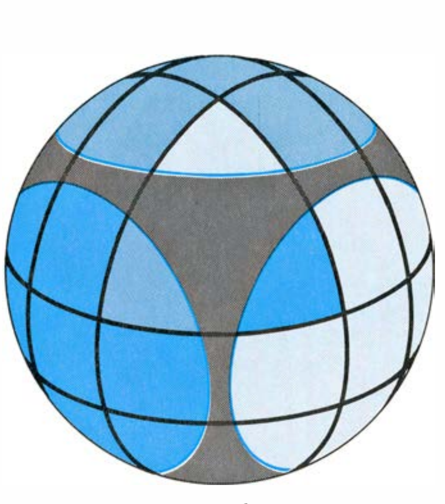
A estas alturas ya basta decir "el Cubo"; nadie necesita aclarar "el cubo de Rubik" para ser comprendido al aludir a este poderoso objeto de intriga y perplejidad. Tanto así, que yo tengo un "cubo" de forma esférica, al que a veces llamo "cubo redondo", pero al que igualmente me refiero diciendo "ese cubo de ahí". Ha sido rebanado en la forma precisa; sus "caras" giratorias y su mecanismo interior son idénticos a los del de Rubik. Y lo que es aún maravilla mayor, tengo también un objeto que pretende pasar por cubo, pero que con absoluta certeza no lo es: se trata de un objeto cúbico seccionado oblicuamente de forma extraña, que se ordena y desordena por diabólicos movimientos al biés [*véanse las ilustraciones a pie de página*]. La esfera es, desde luego, un cubo, mientras que el objeto cúbico es un impostor, travestido de cubo.

Esta proliferación de variedades del "cubo" (término genérico con que a partir de aquí denominaremos estos rompecabezas reordenables por giros) es fenómeno que causa admiración. Erno Rubik (y su eclipsado homólogo japonés Terutoshi Ishige) fueron quienes la iniciaron; después, prendió como el fuego en la pradera. Súbitamente surgieron por doquier variantes sobre el Cubo (el de Rubik, de $3 \times 3 \times 3$, para ser exacto): cubos pequeños, cubos chiquirritines, cubos graciosamente decorados... Ninguno de estos objetos, sin embargo, era sustancialmente diferente del Cubo propiamente dicho.

Los primeros cubos genuinamente diferentes que yo haya visto vinieron del Japón. ¡Eran cubos de $2 \times 2 \times 2$! Uno era magnético, formado por ocho cubitos metálicos que se acoplaban en torno a un imán esférico central. El otro era de plástico, con un intrincado mecanismo análogo, pero no idéntico, al de $3 \times 3 \times 3$ de los cubos de Rubik e Ishige. No puede ser idéntico, porque las piezas maestras del mecanismo de $3 \times 3 \times 3$ son los seis centros de las caras, y en cubos de $2 \times 2 \times 2$ no hay pieza central posible. Más tarde supe que este mecanismo ha sido ideado también por Rubik, y está inspirado en el de $3 \times 3 \times 3$. El cubo $2 \times 2 \times 2$ [*véase la ilustración superior izquierda de la página siguiente*] es un objeto maravilloso e inevitable, en ciertos aspectos más seductor todavía que el de $3 \times 3 \times 3$. Tanto, que me pregunto por qué no los hay en todas partes. El cubo $2 \times 2 \times 2$ (¿cubo de Dubik?) me parece el peldaño ideal para elevarse desde el noviciado más completo a un nivel de cubismo intermedio, pues equivale a resolver solamente los *vértices* del cubo $3 \times 3 \times 3$.

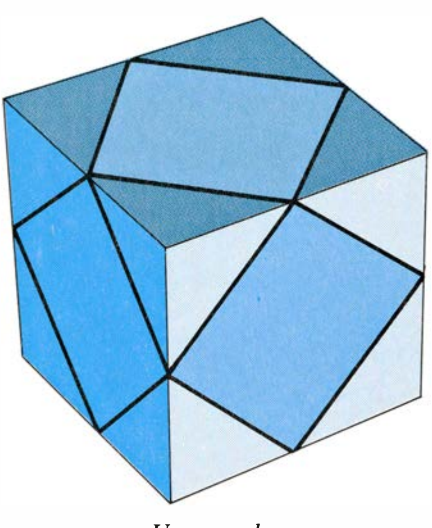
Tal vez no haya debido decir que el cubo $2 \times 2 \times 2$ fue el primero esencialmente diferente que haya encontrado. Mucho antes había visto un "dominó mágico", también invención de Rubik. El dominó viene a consistir en dos de los tres pisos de un cubo $3 \times 3 \times 3$ [*véase la ilustración superior derecha de la página siguiente*]. Sus capas superior y básica, ambas cuadradas, pueden girar de 90 en 90 grados, pero las caras laterales, rectangulares, sólo pueden girar 180 grados cada vez. Otra de las primeras variantes fue el "prisma octogonal", cubo en el que cuatro de sus aristas han sido biseladas, y que al ser retorcido originaba ciertas formas bastante grotescas [*véase la ilustración central de la página siguiente*]. En esta versión se pierde en parte la información sobre paridades de las aristas, pues es imposible decir si las piezas-arista "biseladas" están dirigidas hacia delante o hacia atrás. Por ello, la resolución de esta variante presenta ciertas peculiaridades que la diferencian del cubo completo. En el cubo completo las aristas volteadas se presentan siempre por pares. Otro tanto sucede aquí, pero ahora, al ser imposible discernir si una arista biselada está volteada o no, a veces nos tropezamos con lo que parece ser un cubo resuelto, excepto por una sola pieza de arista, que se encuentra volteada. Si se está habituado al cubo completo esta situación puede resultar desconcertante al principio.

La siguiente variante con que me tropecé fue la ideada por un joven alemán llamado Kersten Meier, quien a la sazón cursaba el segundo ciclo de investigación operativa en Stanford. Meier había construido un prototipo de "pirámide mágica" de funcionamiento bastante burdo. (En realidad, esta pirámide no tiene base cuadrada, como los monumentos egipcios, sino que es un tetraedro regular, cuyas caras son triángulos equiláteros.) El prototipo de Meier era en realidad tan burdo que con frecuencia se desmoronaba al manipularlo. Empero significaba un paso innovador y merecía comercializarse. Más tarde me enteré de que Ben Halpern, matemático de la Universidad de Indiana, había tenido la misma idea. Tanto Meier como Halpern habían ge-





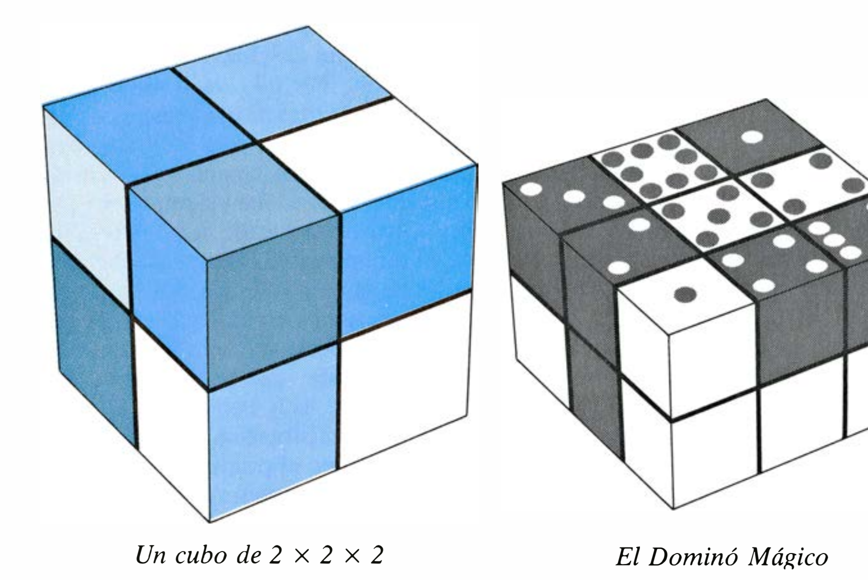
Un cubo



Un no-cubo

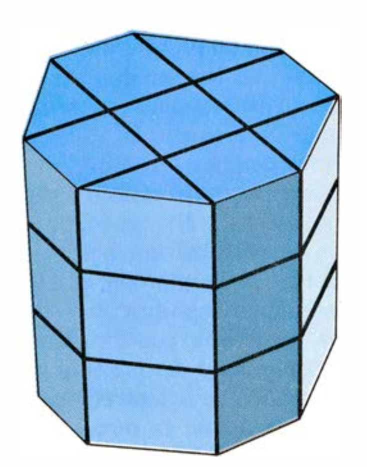
neralizado el mecanismo de Rubik e Ishige, y descubrieron cómo construir un rompecabezas dodecaédrico fundado en iguales principios. Halpern llegó a construir modelos de la pirámide y el dodecaedro [*véanse las ilustraciones inferiores de esta página*].

Al cabo resultó que Uwe Mèffert, otro inventor de origen alemán, venció a Meier y Halpern en el ataque a la pirámide, si bien por diferente procedimiento. En 1972, Mèffert se interesó por las pirámides y por las gratas propiedades que manifiestan al sostenerlas en la mano. De alguna forma dio con la idea de la pirámide de caras giratorias, y construyó algunas [*véase la ilustración de la página siguiente*]. Aunque jugar con ellas era un ejercicio tranquilizante que facilitaba la meditación, tras cierto tiempo las guardó y se olvidó de ellas. Al cabo apareció el cubo de Rubik. Viendo su fenomenal éxito, Mèffert comprendió que su viejo invento pudiera tener valor comercial. Rápidamente patentó su diseño, lo hizo fabricar en cantidad suficiente y se puso en contacto con una compañía juguetera que lo comercializase. El resultado final fue el éxito de la Pyraminx en todo el mundo, un cubo piramidal (cubo en el sentido genérico) de funcionamiento muy diferente del que rige en las pirámides de Meier y Halpern.

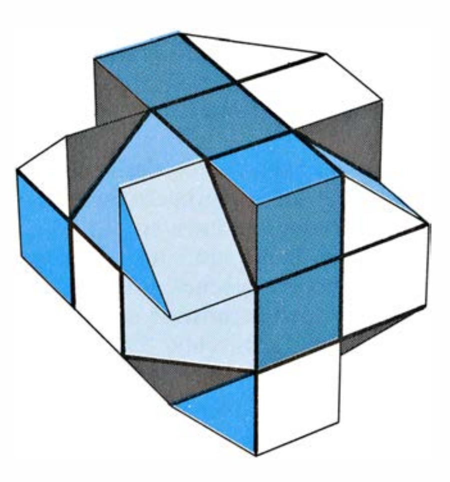


Un cubo de 2 × 2 × 2

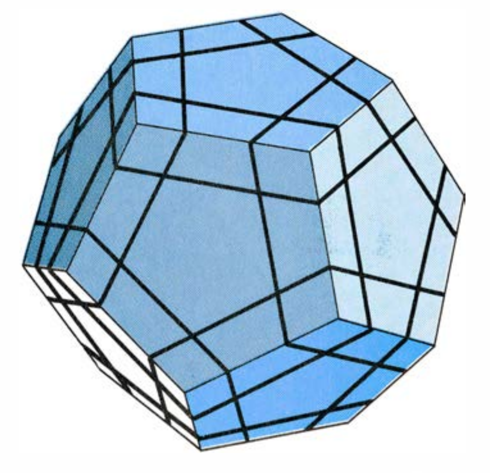
El Dominó Mágico



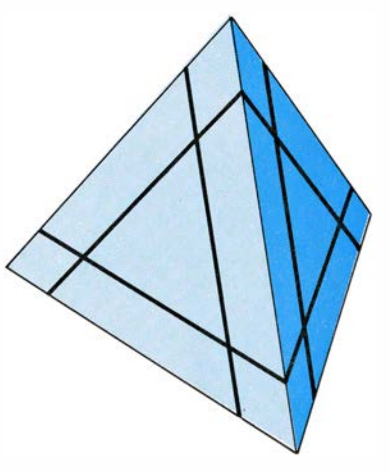
El Prisma Octogonal en estado prístino



Un estado desordenado del Prisma Octogonal



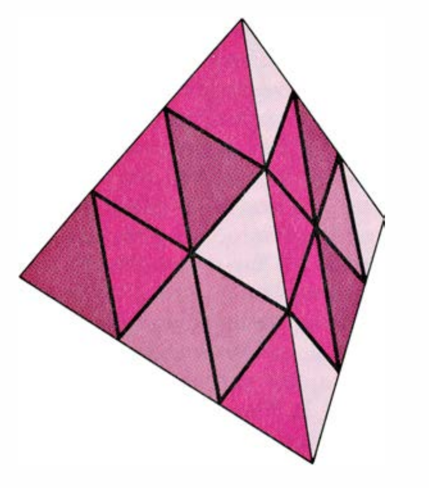
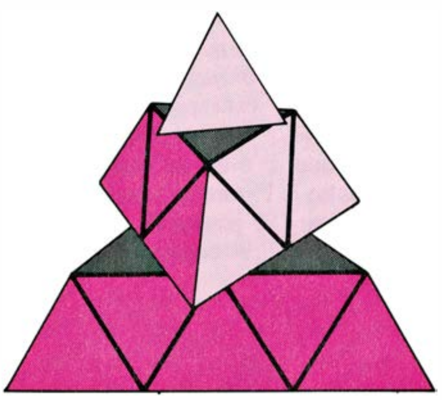
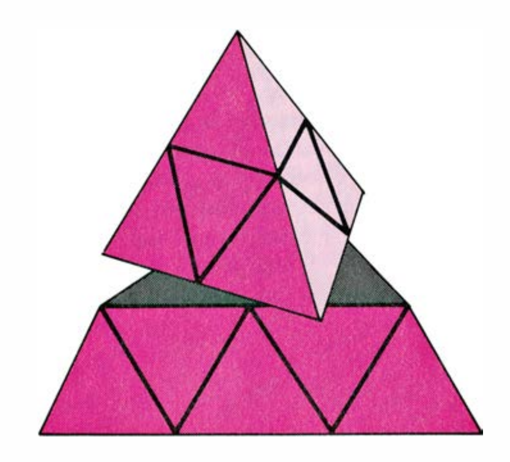
El Dodecaedro Mágico de Pyraminx



El Tetraedro Mágico de Pyraminx

Mèffert, que actualmente reside en Hong Kong, tuvo que dedicarse intensamente a la producción y comercialización de su Pyraminx, y ello le obligó a viajar mucho. Entró así en contacto con otros inventores de todo el mundo y llegó a la conclusión de que sería buena idea lanzar al mercado mundial diversos juguetes de la familia del cubo. Entre esos inventores estuvieron Meier y Halpern. No tardará mucho en que la pirámide de ambos esté en manos de los aficionados. Su nombre será Pyraminx Magic Tetrahedron. También estará disponible el dodecaedro, con la marca Pyraminx Magic Dodecahedron. Si se desea, puede obtenerse un catálogo de la gama completa de productos de Mèffert, escribiéndole a Pricewell (Far East) Ltd., P. O. Box 31008, Causeway Bay, Hong Kong. Mèffert recibirá con agrado ideas para crear rompecabezas nuevos; pero, por favor, no se dirijan con ellas a mí, sino a él. Mèffert proyecta también crear un círculo de aficionados a los rompecabezas, donde los socios abonarían una tasa fija de suscripción recibiendo a cambio un mínimo de seis rompecabezas cada año, que serían ediciones limitadas de formas particularmente complejas o esotéricas de tipo "cúbico". Tendrá mucho gusto en tomar nota de futuros miembros.

Ronald Turner-Smith, amigo de Mèffert y adscrito al departamento de matemáticas de la Universidad china de Hong Kong, es autor de un delicioso librito sobre las configuraciones y la teoría matemática de la Pyraminx, titulado *The Amazing Pyraminx*, que puede adquirirse en rústica por intermedio de Mèffert. En él, Turner-Smith hace para la Pyraminx lo que Singmaster hizo para el cubo en sus *Notes on Rubik's "Magic Cube"*. (Incidentalmente, Singmaster sigue interpretando el papel de oficina de información de la Cubología. Publica un boletín trimestral, titulado *Cubic Circular*. La suscripción cuesta 5

La Pyraminx y sus giros

dólares anuales; escribir al respecto a David Singmaster, Ltd. 66 Mount View Road, London N4 4JR, Inglaterra. Como ya dije, todas las ideas y cuestiones que se les ocurran acerca del cubo deben dirigirlas a Singmaster, no a mí. He de mencionar también que este verano saldrá en Hungría una revista trimestral, *Rubik's*, que podrá adquirirse por 8 dólares anuales a través del Apartado de Correos 223, Budapest 1906, Hungría.) Al igual que Singmaster, Turner-Smith ha puesto a punto un sistema notacional, que utiliza para introducir un poco de la teoría de grupos relacionada con el problema y que permite una comprensión del objeto más profunda que su mera resolución mecánica pueda consentir.

Vale la pena notar que hay dos formas complementarias de manipular y describir la acción de la Pyraminx. Podemos hacer girar bien una cara, bien una pequeña pirámide. Estos dos enfoques son equivalentes y complementarios, pues cada base juntamente con la pequeña pirámide opuesta forman el objeto completo. Turner-Smith considera que las pequeñas pirámides son móviles y las bases opuestas, estáticas. Por mi parte adoptaré ahora este enfoque, retornando más tarde al complementario. Demos pues nombre a los cuatro movimientos posibles [*véase la ilustración superior izquierda de la página siguiente*]. En cada movimiento se puede hacer girar una pirámide vértice, bien la situada arriba (*A*), bien la trasera (*T*), bien la izquierda (*I*) o la derecha (*D*). Las letras *A, T, D, I* denotan giros de 120 grados en sentido horario, mientras que *A', T', D', I'* designan giros de igual amplitud (120 grados), pero en sentido antihorario (mirando hacia el vértice giratorio a lo largo del eje de giro). Observemos que los movimientos dejan en su lugar a los ápices de todos los vértices (aunque los hagan girar). Por consiguiente, podemos considerar que los cuatro vértices son puntos de referencia estáticos, y asignarles papel muy semejante al de los centros de cara en el cubo. Tanto así, que al principio mismo del proceso de resolución podemos hacer girar en un instante todos los vértices para armonizar sus colores, de forma que en lo sucesivo sirvan de recuerdo del color de cada base. Es decir, podemos considerar las cuatro pirámides vértice bien como ornamento decorativo, bien como indicadores útiles.

En el cubo, los objetos elementales que van cambiando de posición suelen llamarse cubitos o cubículos. ¿Cuáles serán aquí los objetos elementales correspondientes? Ciertamente no son las pequeñas pirámides. Al igual que en el Cubo, resulta que existen tres tipos de piezas: piezas-arista, piezas faciales y piezas-vértice o ápices, antes mencionados [*véanse las ilustraciones inferiores de la página siguiente*]. Como podemos ver, en cada vértice hay una pieza facial rodeada de tres triángulos de distintos colores, los mismos que exhibe el apéndice asentado sobre ellos. Al igual que los vértices, esta pieza facial nunca abandona su emplazamiento, pudiendo solamente girar sobre sí misma. En consecuencia, puede decirse que los ápices son partes "trivialmente solubles" de la Pyraminx y que las piezas faciales son partes "fácilmente solubles".

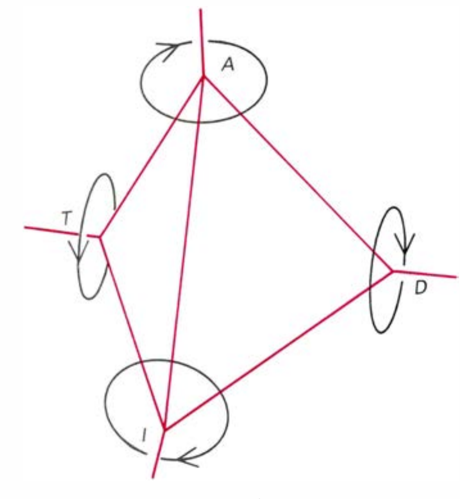
Nos quedan así seis piezas arista, cada una de dos colores; estas piezas pueden viajar por la pirámide, y también, cambiar de orientación, lo mismo que las piezas arista del Cubo. De hecho, resulta que las restricciones que pesan sobre las permutaciones y cambios de orientación de las aristas son exactamente las mismas que las valederas para el Cubo: los cambios de orientación, o "volteos" de las aristas, han de hacerse por pares; las permutaciones de piezas han de ser de clase "par", es decir, permutaciones que impliquen un número par de trueques de posición de las aristas.

Podemos, pues, enumerar rápidamente de cuántas formas distintas podrán estar las aristas distribuidas por la Pyraminx. De no estar sometidas a restricciones, las aristas podrían ocupar los huecos reservados a ellas de 6!, o sea, 720 maneras: la primera arista podría ir a seis sitios, la segunda, a cinco, etcétera. La exigencia de que la clase de tal permutación de aristas sea par obliga a dividir por 2 este número, dando 360 maneras. Por otra parte, si la orientación de las aristas pudiera ser arbitraria, cada una de ellas podría presentarse en cualquiera de sus dos orientaciones, dando 2^6, o sea, 64 posibilidades diferentes. Mas como antes, tenemos que dividir por 2 a causa de las restricciones sobre volteos; disponemos pues de 32 estados de orientación diferentes. Al multiplicar estas dos cantidades resultan 11.520 "estados interesantes" para la Pyramix. Evidentemente, si se tienen en cuenta las posiciones de las piezas faciales y de las piezas vértice, como cada uno de estos grupos tiene 3^4, es decir, 81 modos de girar, sin restricción ninguna, la cifra anterior se infla hasta ¡75.582.720 disposiciones distintas! Quizá lo más verosímil sea descontar las orientaciones de los vértices, pero tomar en cuenta las de las piezas faciales. En tal caso, tendríamos $81 \times 11.520 = 933.120$ estados "no trivialmente distintos" de la Pyraminx.

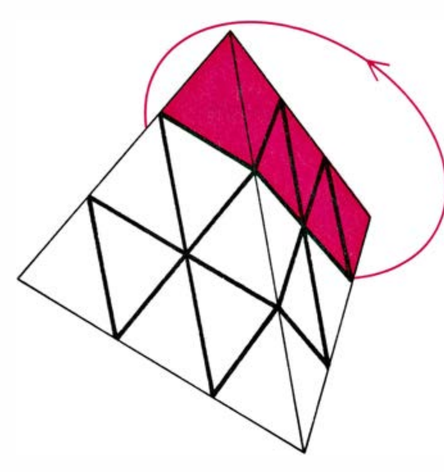
Para resolverla, el más breve de los algoritmos conocidos hasta ahora, descubierto con ayuda de ordenador, requiere 21 giros. Es fácil demostrar que desde ciertas posiciones se requieren al menos 12 giros para retornar al estado inicial; no se conocen, sin embargo, ni la naturaleza del "algoritmo de Dios", que toma siempre la ruta más breve po-

sible, ni el número máximo de giros que pueda exigir, como tampoco se conocen para el Cubo.

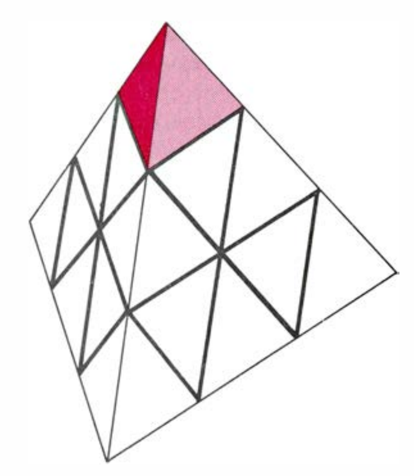
Cuando Mèffert ideó la Pyraminx tenía clara conciencia de que su interior admitía otras soluciones, aunque mantuviera idéntico su aspecto exterior, con nueve triangulitos por cara. Por consiguiente, ideó él mismo otros posibles mecanismos interiores que permitieran más rica variedad de giros. El objeto que acabo de describir es la llamada Popular Pyraminx. La Master Pyraminx es una versión diferente, cuya salida al mercado está prevista para principios del año próximo. En ella, además de poderse ejecutar todos los movimientos de la Pyraminx, cada arista puede pivotar 180 grados en torno a su punto medio, lo que consiente intercambiar dos piezas vértice cualesquiera al tiempo que se voltea solamente una de las aristas [*véase la ilustración superior derecha de esta misma página*]. Tal flexibilidad obliga a que cada pieza facial se descomponga a su vez en varias subpiezas, algunas de las cuales pueden viajar en torno a toda la pirámide.



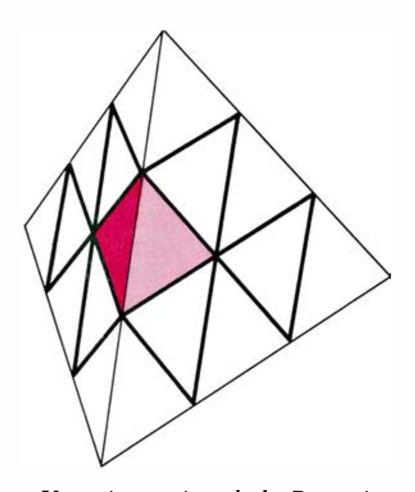
Notación torsional de la Pyraminx



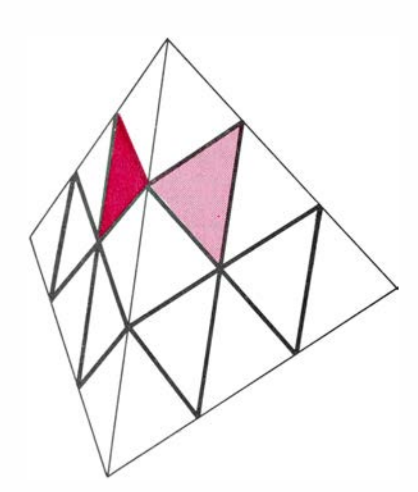
Modo secundario de torsión de la Senior Pyraminx



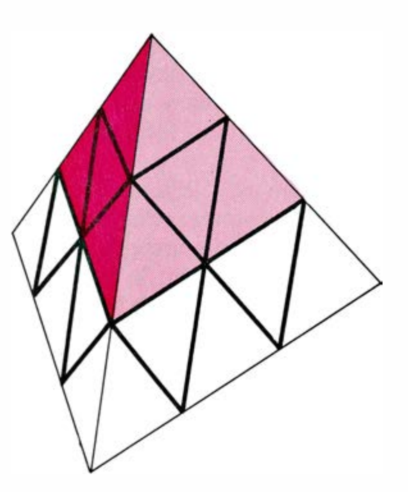
Un ápice de la Pyraminx



Una pieza arista de la Pyraminx



Una pieza facial de la Pyraminx



Una pirámide menor

Tenemos así un rompecabezas mucho más complicado. Su mecanismo es superlativamente ingenioso, porque en el proceso de pivotado cada uno de los vértices viajeros permanece conectado al resto de la Pyraminx gracias a una piececilla interior invisible situada en el bloque facial ahora fragmentado. Esta piececilla ignora a qué arista debe vasallaje, y de no ser por un hábil recurso ingenieril, la piececilla invisible y el vértice se desprenderían. El ingenioso mecanismo permite que cada vértice "enclave" su piececilla en la arista adecuada antes de comenzar el pivotado, y la "desenclave" al terminar éste. Según Turner-Smith, el número de estados desordenados de la Master Pyraminx rebasa los 446 billones.

Mordido ya por el gusanillo del cubo, Mèffert no se detuvo aquí, adentrándose más todavía en el mundo de los poliedros regulares. Su paso siguiente consistió en diseñar un octaedro de ocho colores, cada una de cuyas caras triangulares está descompuesta en nueve triangulitos. ¿De qué modos podremos hacerlo girar? Al igual que con la Pyraminx, Mèffert se percató de las distintas posibilidades de giro y torsión. Debe observarse que las dos posibles descripciones de los movimientos de la Popular Pyraminx dejan de ser equivalentes en el octaedro. Recordemos que antes se trataba de hacer girar, bien caras, bien pequeñas pirámides. La razón de que ambas descripciones fuesen que el giro de una cara es complementario del giro de la pequeña pirámide tario del giro de la pequeña pirámide opuesta. En los octaedros, en cambio, un giro facial de 120 grados no es complementario de la rotación de 90 grados que pueda darse a ninguna de las pequeñas pirámides centradas en un vértice [*véase la ilustración superior de la página siguiente*]. Comprendiendo que así disponía de un grado de libertad mayor, Mèffert ideó sendos mecanismos para cada una de las dos posibles formas de giro.

El octaedro que pronto saldrá al mercado (con la nada imaginativa denominación de Pyraminx Magic Octahedron) será el modelo donde las seis pirámides menores son giratorias. Hay en ella, por tanto, tres ejes de giro ortogonales, como en el Cubo. Este paren-

tesco, de apariencia trivial, comporta en realidad mucho más que un grano de verdad. En efecto, el octaedro de Mèffert y el Cubo de Rubik son afloramientos en superficie de una misma idea profunda y abstracta. Para comprenderla, observemos que un cubo y un octaedro son sólidos *duales* uno de otro: los centros de las caras de una de las figuras son vértices de la otra. Así pues, los seis centros faciales de un cubo definen un octaedro, y los ocho centros faciales de un octaedro son los vértices de un cubo.

Imaginemos un Cubo, e inscrito en su interior el octaedro definido por los centros de las caras de aquél [*véase la ilustración inferior izquierda de esta página*]. Cada torsión de una cara del Cubo induce una torsión de la correspondiente pirámide del octaedro. Parecería a primera vista que cada uno de los estados desordenados del Cubo equivale a un estado desordenado del octaedro. Sin embargo, esto no es enteramente cierto. Para comprenderlo bien es necesario ver qué pieza se aplica sobre cuál en la correspondencia Cubo-octaedro. Lo mismo que la Popular Pyraminx, el octaedro tiene vértices, piezas faciales y piezas de arista. Como antes, los vértices apenas si tienen más papel que el ornamental, y las piezas faciales giran como un todo. Por tanto, las piezas faciales del octaedro, más sus decorativos ápices, se corresponden con los centros de cara del Cubo. Nos quedan así en el octaedro las piezas arista, y resulta claro que éstas, por ser diédricas, tienen que corresponderse con las piezas arista del Cubo. ¿Y qué ha sucedido con los vértices del Cubo? ¿Con qué se corresponden? Con nada. Los vértices del Cubo carecen de homólogos en el octaedro, que es, por consiguiente, simplificación de aquél.

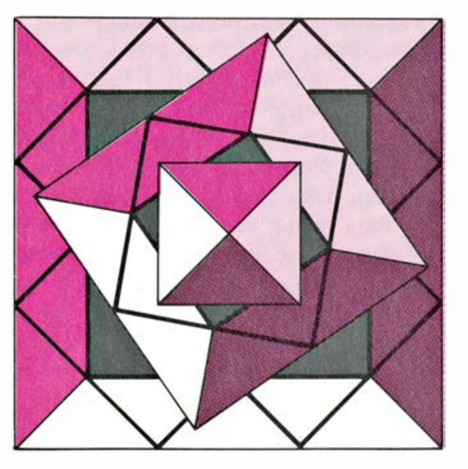
Para visualizar adecuadamente la correspondencia Cubo-octaedro debemos colorear alternativamente uno de estos rompecabezas. Puesto que estamos más familiarizados con el Cubo, veamos cómo modificarlo para "convertirlo" en octaedro mágico. El coloreado correcto no se centra en las caras, sino en los vértices [*véase la ilustración inferior derecha de esta página*]. Stan Isaacs, especialista en ciencias de cómputo y excelente aficionado a los rompecabezas, ha trucado uno de sus cubos mágicos (tiene docenas) para que simule un octaedro de Mèffert.

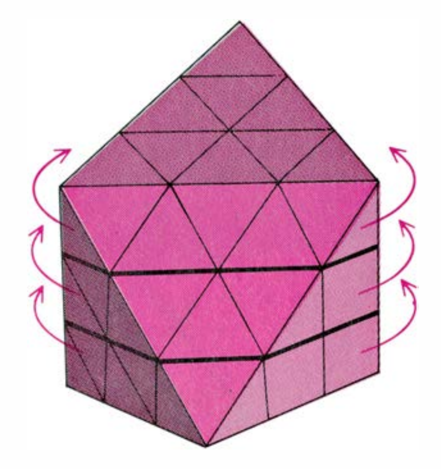
Quienes hayan adquirido soltura en resolver el Cubo de $3 \times 3 \times 3$ coloreado de la forma habitual, descubrirán que su habilidad con él no basta para manejar el cubo extrañamente coloreado de Isaacs, porque ahora sí hay que tener en cuenta la orientación de los centros de cara. Por otra parte, se tiene la correspondiente simplificación: en este cubo ya no existen "quarks". Es decir, no pueden existir cosas tales como vértices con torsión, y ni siquiera vértices mal situados, porque todos los cubitos triédricos son ahora blancos por sus tres caras. En el Cubo ordinario, los quarks (vértices con torsión $+ 1/3$, es decir, 120 grados en sentido horario) y los antiquarks (cuya torsión es $- 1/3$, o sea, 120 grados en sentido antihorario) dan siempre, al sumarlos, una torsión total de valor entero.

Todo cuanto se precisa para resolver este cubo (o el octaedro) es la habilidad de devolver a su sitio las aristas y los centros de las caras (aunque ahora, con la novedad de respetar las orientaciones). Evidentemente, no todo octaedro mágico será equivalente a un simple recoloreado del Cubo $3 \times 3 \times 3$, porque sus ejes de giro pueden ser otros. En particular, el otro modo de torsión considerado por Mèffert (en el cual las caras del octaedro giran 120 grados) no tiene relación alguna con el Cubo.

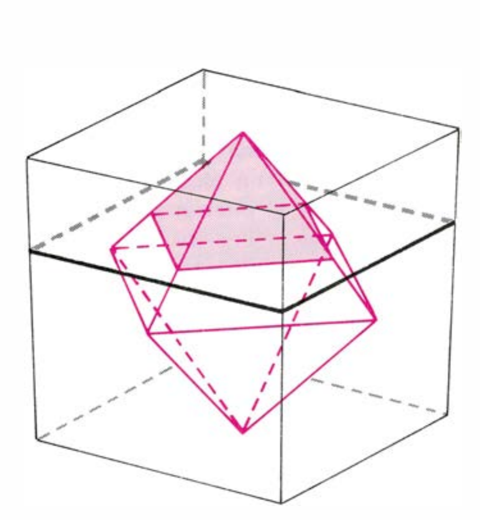
En su catálogo de 1982, Mèffert presenta una figura de icosaedro (¡adivine el lector con qué nombre!) cuyas 20 caras triangulares no están descompuestas; valsean de cinco en cinco, en torno a los 12 vértices [*véase la ilustración superior de la página siguiente*]. Puesto que el movimiento está centrado en los vértices y no en las caras, inmediatamente se nos ocurre pensar en el sólido dual, el dodecaedro. El rompecabezas dual tendría entonces movimientos centrados en caras, análogamente a como el dual del octaedro, de giros centrados en los vértices, es el Cubo, de giros centrados en las caras. (Incidentalmente, ¿cuál será el rompecabezas dual de la Pyraminx?)

En realidad, el catálogo de Mèffert presenta también otros dos rompecabezas dodecaédricos. De ellos, el menos complicado es el llamado Pyraminx Ball (que abreviaremos Bola); el otro, de bellísima filigrama, es el Pyraminx Crystal (brevemente, Cristal) [*véanse*

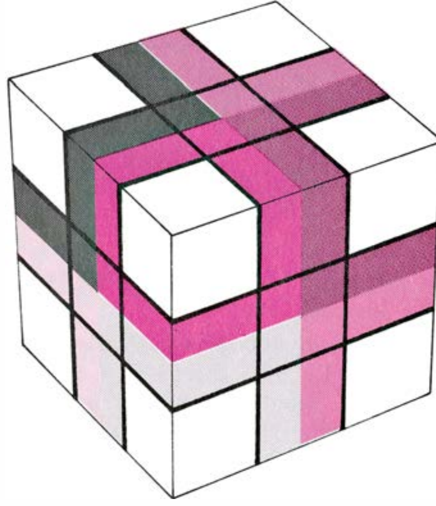




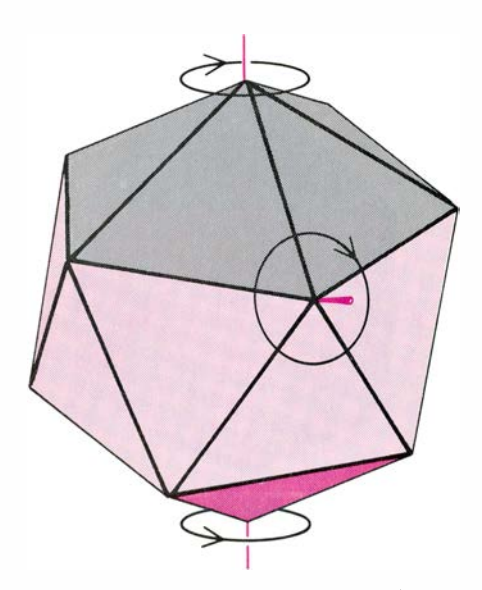
*Octaedro de Mèffert (***izquierda***) y su otra posible torsión*



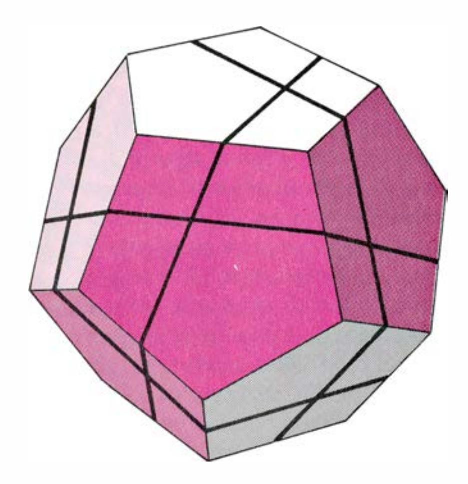
Los giros del Cubo y el Octaedro están relacionados



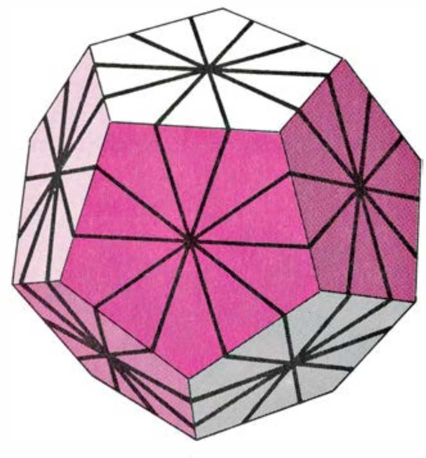
El cubo de Isaacs remeda al Octaedro de Mèffert



Icosaedro Mágico de Pyraminx



La Bola de Pyraminx



El Cristal Mágico de Pyraminx

las ilustraciones central e inferior de esta página]. La Bola tiene cuatro ejes de rotación, igual que la Pyraminx; el Cristal tiene seis. Su aparición en el mercado es inminente.

Llegados a este punto el lector se preguntará si podrá existir un cubo –y me refiero con ello a un cubo genuíno, hexaédrico, de caras cuadradas– con un mecanismo de giros centrados en los vértices. ¡Dicho y hecho! Esto es, en efecto, el objeto cúbico que puede torcerse al biés reproducido en la ilustración derecha de la página 103 y en la superior izquierda de la página siguiente. Para verlo durante el proceso de torsión, fíjese en la figura superior derecha de esa ilustración. Un periodista inglés, Tony Durham, fue el primero que tuvo esta idea. Durham mostró a Mèffert su diseño, quien lo desarrolló y preparó para su lanzamiento comercial, dotándolo además de ciertos detalles que habían probado ser útiles en la Pyraminx. A mí me gustaría bautizarlo "Cubiés" o tal vez, "Sesgubo", pero Mèffert ha optado por el más prosaico nombre de Pyraminx Cube.

Cada uno de los cuatro tajos dados al Sesgubo lo escinde en dos mitades iguales. Cada sección es plano bisector de una de las cuatro diagonales espaciales del cubo. Si el lector se para a pensar un momento se dará cuenta de que la traza producida por la sección sobre las caras define un hexágono regular que circunda al cubo. Cada tajo pasa por las seis caras, y por eso, cada giro afecta simultáneamente a las seis. En este aspecto, el Sesgubo es más perverso que el Cubo, donde los giros dejan siempre una cara intacta. A pesar de la simplicidad de este objeto, cuesta mucho habituarse a sus giros al biés. Aunque, claro, ello forma parte de su atractivo.

Durham nos ofrece comentarios francamente penetrantes acerca de su invento en una serie de notas escritas por él y tituladas "Rompecabezas de cuatro ejes". Me tomaré la libertad de espigar de ellas algunos párrafos:

"El grupo de simetrías generado por cuatro ejes ternarios es el grupo de giros del tetraedro, y tiene orden 12. Casi todos los poliedros más conocidos, tanto regulares como semirregulares, poseen esta simetría tetraédrica, si bien el grupo de sus simetrías propias puede ser mucho más rico. Por consiguiente, podemos situar un mecanismo de eje cuádruple en el interior de cualquier rompecabezas poliédrico de forma regular o semirregular, y el objeto conservará su forma durante el juego. La Bola puede parecer extraña a primera vista, pero sin duda pone de relieve cómo las simetrías del tetraedro subyacen a las mucho más variadas del dodecaedro.

"El mecanismo del Cubo de Rubik no tiene esta propiedad. Sus ejes de rotación son ejes cuaternarios (cuatro giros de 90 grados alrededor de cada uno), y estos ejes se encuentran, por lo común, solamente en la familia de sólidos cubo-octaédricos. Puede "construirse" un dodecaedro a partir de un cubo de Rubik; mas para preservar su forma durante los giros, es preciso que éstos sean de media vuelta. Los giros de un cuarto de vuelta invocan simetrías que el dodecaedro no posee.

"Todos los rompecabezas de eje cuádruple tienen una bola o husillo central. Cuatro piezas (vértices, por lo general) están directamente sujetas a la bola. La Pyraminx ordinaria tiene seis piezas libremente "flotantes", provistas de pestañas que enganchan por debajo de las piezas vértice. En el Pyraminx Cube, las piezas libremente flotantes análogas son los centros de las caras, y son cuadradas. En la Bola dodecaédrica, las piezas de cuatro facetas desempeñan idéntico papel.

"El Pyrminx Cube y la Bola tienen cuatro piezas flotantes más, vértices lo mismo que antes. Estas piezas poseen sus propias pestañas, que en la posición inicial se alojan por debajo del primer grupo de piezas flotantes. Hay por tanto tres niveles jerárquicos de piezas mutuamente trabadas, conceptualmente semejantes a las del sistema de Rubik, pero geométricamente muy distintas.

"Las ocho piezas triédricas del Pyraminx Cube ofrecen idéntico aspecto. A primera vista pudiera pensarse que dos cualesquiera son permutables de lugar. En realidad hay cuatro esquinas que flotan libremente, mientras otras cuatro están rígidamente fijadas a la bola central. Estos dos tipos nunca podrán intercambiar sus papeles. Asimismo, la forma cuadrada de las piezas centrales de las caras es engañosa; por el interior, las partes mecánicas de estas piezas no son ya tan simétricas. Tales piezas no podrán jamás retornar a sus posiciones iniciales (con respecto al sistema rígido definido por los cuatro vértices) al ser girada 90 grados. Solamente son posibles los giros de media vuelta.

"La Pyraminx ordinaria tiene cuatro puntos fijos evidentes, a saber, sus cuatro vértices. Comparada con un Pyraminx Cube, y sabiendo que cuatro vértices de éste son fijos y cuatro flotantes, uno se pregunta cuáles corresponden a cuáles. En realidad, no hay diferencia. Los cuatro vértices libres se mueven independientemente de los fijos, si bien al moverse lo hacen conjuntamente, como si estuvieran físicamente ligados."

A continuación, Durham da para la Pyraminx la notación *A, T, I, D* de Turner-Smith, mencionando que tal

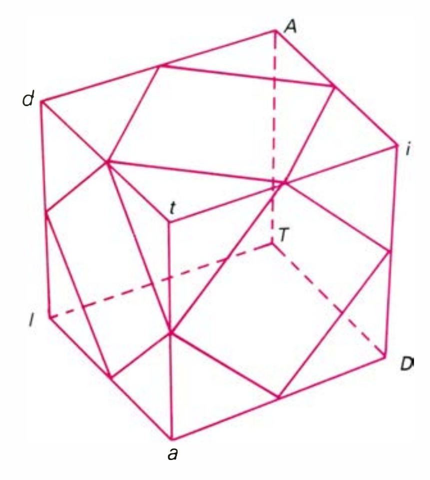
notación puede adaptarse a cualquier rompecabezas de eje cuádruple (como por ejemplo su Sesgubo), sin más que nombrar *A, T, I, D,* a cuatro centros de giro. (En la Pyraminx podríamos denotar así, ya los cuatro vértices, ya los cuatro centros faciales. En el Sesgubo serían cuatro de los ocho vértices; los otros cuatro quedarían sin bautizar.) Podemos ahora transcribir cualquier movimiento. Si su centro está en uno de los vértices con nombre, bastará darle el mismo. Si el movimiento está centrado en uno de los vértices anónimos se usará el nombre del movimiento complementario, ya que no importa qué mitad del rompecabezas sea la girada. (Quizás el lector tenga que pensárselo un poco; aunque parece un punto difícil, en realidd es trivial.) Durham señala que a veces conviene disponer de nombre para los cuatro lugares restantes y para los giros en torno a ellos. A tal fin se vale de las letras *a, t, i, d*. Por tanto *A* y *a* producen internamente iguales efectos en el rompecabezas, si bien al alzarlo en el aire nosotros vemos que la orientación general es diferente. Y aunque Durham concede que la notación compuesta se presta a confusiones, en ocasiones aboga por ella.

"A veces es preciso combinar ambas notaciones para mejor comprender lo que está ocurriendo. Así, *A t A' t'* es un elemento de una útil clase de transformaciones, los llamados 'conmutadores' (dos movimientos seguidos de sus inversos, y por tanto, de la forma general *x y x' y'*) pero tal vez nunca sospechásemos que así es al describirlo en notación de coordenadas regulares (*A T I' T'*) o alternativas (*a i a' t'*).

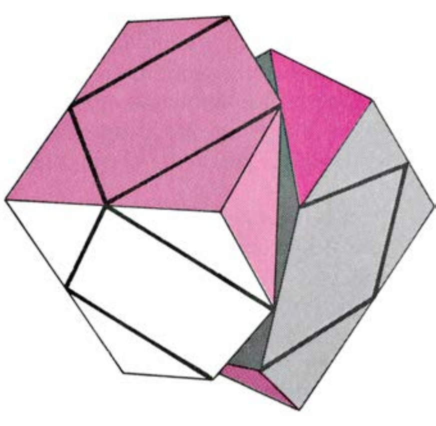
"El Pyraminx Cube y la Bolsa pueden ser descritos como rompecabezas de 'corte profundo', en contraste con los rompecabezas de 'corte somero', como el Cubo de Rubik. En estos últimos, los cortes se dan cerca de la superficie; en los rompecabezas de corte profundo el tajo pasa cerca del corazón del objeto. En los de corte somero, durante los giros el 'cuerpo principal' permanece estacionario al tiempo que se voltea una parte pequeña del rompecabezas. En los de sección profunda dudamos de cuál ha sido la porción girada y de cuál ha permanecido estacionaria. Por eso, al manejarlos es preciso tomar seriamente en cuenta los sistemas alternativos de coordenadas.

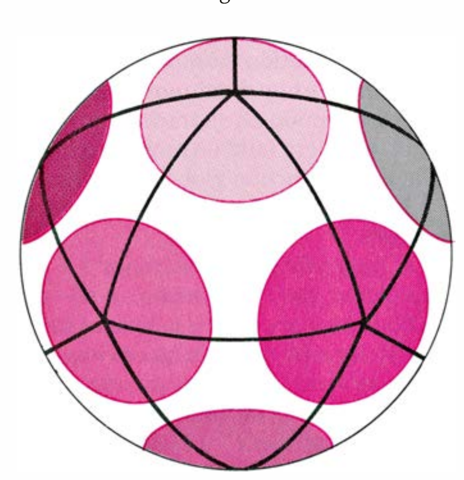
"Los rompecabezas de sección profunda dictan también un enfoque 'global' de la solución. En ellos resulta particularmente difícil trabajar en una zona sin afectar también a las demás. Sin embargo, conforme se avanza hacia la solución este hecho mismo viene en nuestra ayuda. La torsión de los vértices se va esfumando en mágica sincronía. El último volteo, el último trueque de posiciones se nos da hecho automáticamente. Conforme nos acercamos para rematarlo, miles de millones de vías de escape por donde el rompecabezas podría fugársenos se van cerrando. Entran en funcionamiento las ligaduras de paridad; y cuando cada movimiento activa cinco u ocho ciclos de permutaciones, como sucede en los juegos de sección profunda, las ligaduras de paridad tienen gran fuerza."

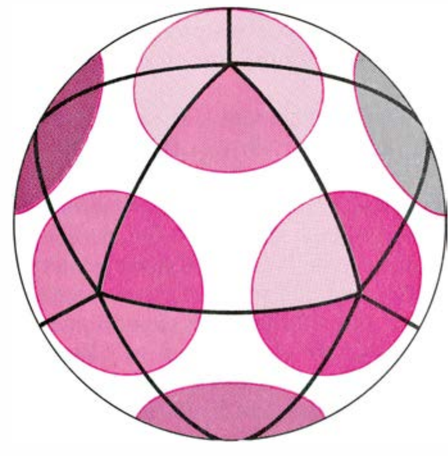
En la sección donde las notas de Durham tratan de las restricciones de paridad, hace constar la siguiente disculpa: "Ruego perdón por lo relajado de la forma en que empleo el término

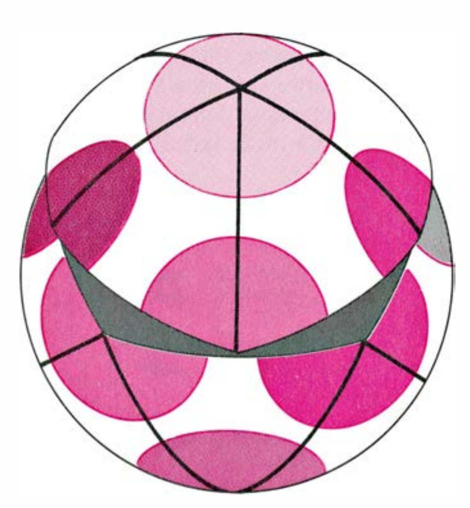


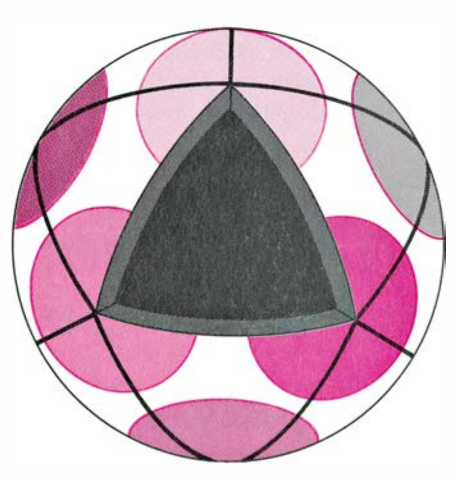
Notación de los ocho vértices del Sesgubo


El Sesgubo a mitad de una torsión


Estado prístino de la IncreíBola


Un quark visible sobre la IncreíBola


"Accidentada" torsión de la IncreíBola


La IncreíBola, sin un icosalito

'paridad', que no sólo da cabida a condiciones de divisibilidad por 3 (y no estrictamente por 2), sino que invoca también nociones aún más distantes. Emplearemos el término 'restricción de paridad' para referirnos a cualesquiera restricciones que por pesar sobre transformaciones imaginables del rompecabezas impidan su realización en funcionamiento normal. La lista no contiene, por ejemplo, el mandamiento 'No desearás trocar piezas faciales con vértices'. Sería demasiado traído por los pelos. Pues si admitimos como imaginable tal transformación, igualmente podríamos imaginar movimientos que conviertan el rompecabezas completo en uranio agotado, o en un queso manchego."

A continuación, Durham enumera todas las restricciones de "paridad" del Sesgubo, en el sentido generalizado que da al término:

"1. Los cuatro vértices (fijos) *ATID* pueden permutarse entre sí, y lo mismo los cuatro vértices restantes, *atid*; en cambio está prohibido mezclar los de uno y otro tipo.

2. Los *ATID* se mueven como una unidad tetraédrica rígida. Esta restricción sólo afecta a sus posiciones relativas en el espacio, no a sus orientaciones.

"2*a*. Exactamente por iguales razones, los cuatro vértices (libres) *atid* se mueven también como una unidad tetraédrica. Además, su movimiento es independiente de *ATID*. En realidad, cualquiera de las 12 posiciones relativas posibles de *atid* y de *ATID* es alcanzable en dos movimientos cuando más.

"Si bien los *ATID* son fijos y los *atid* son libremente flotantes, matemáticamente la restricción 2 y la restricción 2*a* tienen exactamente igual categoría. Al tratar del Cubo de Rubik, los autores han considerado 'inimaginable' la transposición de dos centros faciales, mientras que la transposición de dos piezas diédricas 'está prohibida, pero es imaginable'. Por analogía con este convenio, la restricción 2*a* cuenta como restricción de paridad, mientras que 2 no puede ser considerada como tal. Como esta situación es muy poco satisfactoria, resulta urgente una definición de 'paridad' más rigurosa. ¿Es cuestión de geometría? ¿De mecánica? ¿De topología? Adviértase que el problema consiste en enumerar las posiciones *imposibles*; las *posibles* se cuentan sin dificultad.

"3. La suma de torsiones de los vértices *ATID* es siempre igual, módulo 3, al estado de torsión del rompecabezas, tomado globalmente.

"(Aquí, 'torsión' se refiere solamente a los vértices, y su valor es 0, +1 o −1. La torsión de un vértice se mide con relación al tetraedro rígido del que forma parte. Por tanto, la torsión de *A* se mide con relación a *ATID*. Las torsiones en sentido horario de los vértices cuentan como +1, y las de sentido antihorario, como −1. Por el contrario, el retorcimiento o torsión global del rompecabezas es únicamente función de las *posiciones* de los vértices, y no de sus orientaciones. Si las posiciones relativas de *ATID* y *atid* son las correspondientes a la posición inicial, el 'retorcimiento' es nulo. Si para retornar a la posición de partida basta un solo movimiento de sentido horario, la torsión global será −1, y si tal movimiento ha de tener sentido antihorario, entonces +1. Y si hiciera falta un movimiento de cada tipo, la torsión global sería nula.)

"3*a*. Repetición de 3, pero ahora con *atid*.

"A partir de las restricciones 3 y 3*a* deducimos que la torsión total de *ATID* es siempre igual a la torsión total de *atid*. Resulta también que es imposible hacer girar 120 grados un sólo vértice (esto es, crear un quark aislado.) Podríamos parafrasear 3 y 3*a* diciendo que el rompecabezas 'sabe' de tres maneras distintas a cuántos giros se encuentra (módulo 3) de la posición inicial.

"4. Es imposible transponer exactamente dos piezas faciales.

"5. Es imposible que ninguna pieza facial gire sobre sí misma 90 grados.

"6. Es imposible voltear en 180 grados a una sola pieza facial."

+ 0 0 0 − 0 0 0	π^+ mesón ($u\bar{d}$)	+ + 0 0 + 0 0 0	protón (*uud*)	+ 0 0 0 + + 0 0	neutrón (*udd*)
+ + + 0 0 0 0 0	Δ^{++} (*uuu*)	0 0 0 0 + + + 0	Δ^- (*ddd*)	− − 0 0 − 0 0 0	antiprotón ($\bar{u}\bar{u}\bar{d}$)
− 0 0 0 + 0 0 0	π^- mesón ($\bar{u}d$)	+ − 0 0 0 0 0 0	η^0 mesón ($u\bar{u}$)	0 0 0 0 + − 0 0	π^0 mesón ($d\bar{d}$)

"Hadrones" compuestos por quarks del Cubo

Durham da demostraciones de estas interesantes propiedades, pero como en su mayor parte son semejantes a las correspondientes al Cubo, las omitiré. Teniendo en cuenta todas las restricciones anteriores, Durham calcula el número total de estados desordenados de su Sesgubo, que son 100.776.960. Tal supone, sin embargo, que se tiene forma de saber en qué orientación se encuentran las caras, lo cual no sucede, salvo que las marquemos nosotros. Por tanto, el número de estados *visualmente distinguibles* se reduce en cinco factores 2, quedando finalmente en 3.149.280. Se trata de una cifra bastante menor que la del Cubo (4×10^{19}), pero ciertamente la dificultad no disminuye en proporción al número de estados. (¿Cabe siquiera imaginar cómo sería un rompecabezas 10 billones de veces más fácil que el Cubo de Rubik?)

Las observaciones finales de Durham elevan a un nivel más excelso todavía el precioso paralelismo que Solomon W. Golomb estableció entre fenómenos cubológicos y partículas elementales. Golomb hizo notar que muchas partículas elementales tienen correspondientes en el Cubo de 3 × 3 × 3. Entre ellas se encuentran los quarks (q), los antiquarks ($\bar{q}$), los mesones (pares $q\bar{q}$), los bariones y antibariones (tríos qqq y $\bar{q}\bar{q}\bar{q}$). Durham generaliza esta analogía como sigue:

"La definición de 'torsión' ha de modificarse con objeto de acomodarla a la física de partículas. A los giros de sentido horario aplicados a uno de los vértices *ATID* se les asignará ahora el valor + 1/3, y lo mismo a cualquier torsión *antihoraria* de alguno de los vértices *atid*. Cada uno de ellos será un *quark*. Sus opuestos son los antiquarks, de valor −1/3. Como se verá, la torsión corresponde al *número bariónico*. La torsión total de los vértices es siempre un entero. Los movimientos individuales del rompecabezas son siempre mesones.

"Los quarks situados en los vértices *ATID* se considerarían 'arriba' (*up*), abreviadamente, quarks u; los situados en *atid* son 'abajo' (*down*), quarks d. Ambos quarks tienen espín isotópico de valor 1/2, y se los distingue por la orientación del vector isospín en un espacio abstracto. La proyección I_z del isospín tiene valor +1/2 para el quark u, y valor −1/2 para el quark d. De no presentarse propiedades de extrañeza, encanto, etcétera, la carga eléctrica de una partícula viene dada por $Q = I_z + B/2$, donde B es el número bariónico. Por tanto, los quarks u tienen carga elétrica de 2/3, mientras los quarks d

tienen carga −1/3. (Para los antiquarks basta multiplicar por −1 todos los números cuánticos.) Y nuevamente el rompecabezas sirve de modelo para una importante característica de la realidad observada, a saber, que la carga eléctrica de todas las partículas ha de ser un número entero.

"Los números cuánticos significativos de nuestros dos quarks son como sigue:

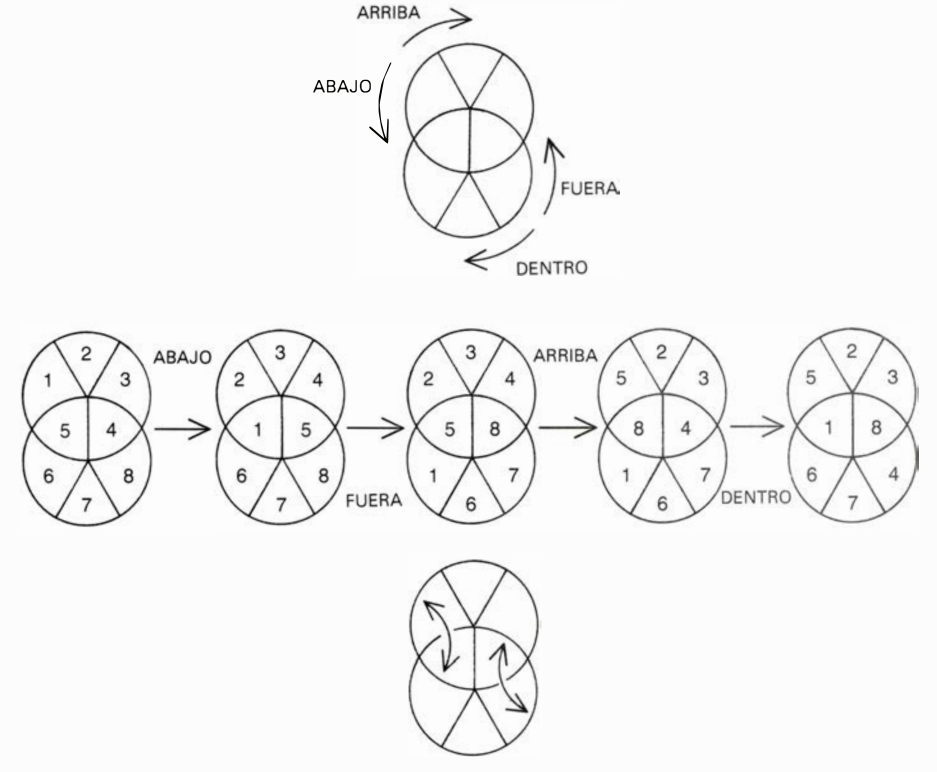
	u	d
B	1/3	1/3
I	1/2	1/2
I_z	1/2	−1/2
Q	2/3	−1/3

"Podemos ahora ensamblar varios 'hadrones', partículas que interactúan fuertemente [*véase la ilustración de la página anterior*]. Cada partícula está representada por dos hileras, cada una con cuatro símbolos. Los cuatro lugares de la hilera superior representan las torsiones de los vértices *ATID*; en la inferior se hace otro tanto para los vértices *atid*. Los quarks están denotados '+' y, los antiquarks, '−'.

"La simetría isotópica es una simetría global, y la fuerza (nuclear) fuerte es invariante frente a transformaciones que hagan girar el vector espín isotópico en igual magnitud para todas las partículas. Semejante transformación convertiría todos los quarks u en quarks d y viceversa, y ello sin discontinuidad. Protones y neutrones intercambiarían papeles. Para el rompecabezas, el proceso correspondiente sería una *rotación continua del objeto completo en el espacio*. En efecto, puede llevar los vértices *ATID* a la posición que antes ocupaban los *atid*, con lo que un quark 'arriba' se convierte en un quark 'abajo'.

"Todo esto carece de importancia en lo concerniente a la 'interacción fuerte' (es decir, al funcionamiento normal del rompecabezas). Los vértices *ATID* y *atid* son funcionalmente idénticos. Pero sí cuenta cuando se intenta desmantelar el objeto, al descubrirse entonces que un grupo de vértices está anclado al núcleo del juguete, y el otro grupo, no. Tales operaciones de desmontaje pueden considerarse 'interacciones débiles' o electromagnéticas, y pueden violar las reglas de conservación a que obedece la interacción fuerte. En realidad, violan demasiado bien tales reglas, pues permiten la creación de quarks individuales aislados."

Durham hace notar que la analogía presenta todavía puntos flacos: no hay conservación de carga o del número bariónico; no está definido el análogo del espín; tan sólo se han podido representar dos "sabores" de quark (arriba y abajo); el "color" de los quarks no tiene traducción al rompecabezas... Entretanto, Golomb ha estado buscando activamente un procedimiento para representar el color de los quarks en la analogía por él ideada con el Cubo 3 × 3 × 3.

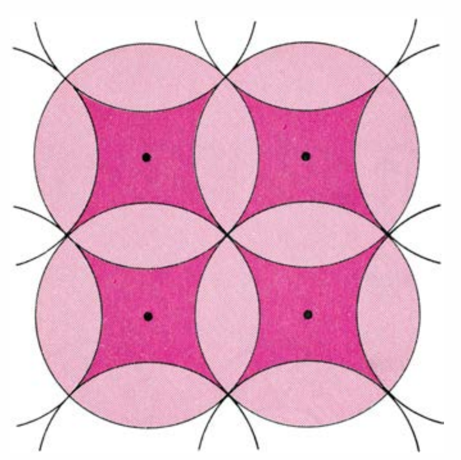


Transformación del conmutador "abajo-fuera-arriba-dentro" (ba-fu-ri-den). Vemos abajo el resultado

Sean cuales fueren las carencias de esta analogía, se trata de una de las más llamativas, a mi juicio, que se haya podido descubrir jamás. Mucho me asombraría que fuera pura coincidencia. Por algún motivo, no puedo sino estar convencido de que las fascinantes pautas comunes a estos rompecabezas macroscópicos y a las partículas subatómicas son manifestación de un cierto orden subyacente, de cierto sistema de principios comunes a ambos. Tanto así, que tengo fe en que, contemplados con enfoque apropiado, los principios de teoría de grupos que gobiernan las restricciones de paridad de los "cubos" podrán transferirse al dominio de la física subatómica, suscitando ideas nuevas acerca de las simetrías observadas en ella. ¡Si este argumento no basta para aguijonear a alguno de los físicos que estén leyendo el artículo, no sé de nada que pueda lograrlo!

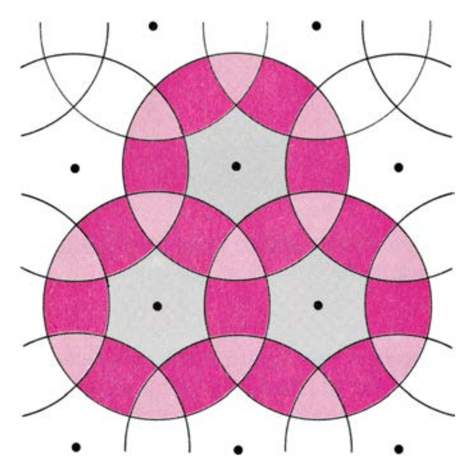
Quizá mi "cubo" favorito sea el que he apodado IncreíBola [*véase la ilustración central izquierda de la página 109*], que ha sido ideada por un maestro alemán de Dortmund, llamado Wolfgang Küppers y que figura también en el catálogo de Mèffert. En el momento de escribir esto tal vez sea yo campeón del mundo de IncreíBola (si no del mundo, tal vez sí de mi manzana), cuya resolución me exige por término medio unos cinco minutos. Estoy seguro, empero, de que mi gloria no será duradera en cuanto el rompecabezas sea comercializado este verano por la Milton Bradley Company. Su nombre comercial americano será Impossi*Ball.

La I-Bola es en esencia un dodecaedro redondeado, cada una de cuyas doce caras (dodecalitos, los llamo yo) ha sido dividido en cinco triangulitos elementales. Hay así 60 de estos triangulitos. Si en lugar de grupos de cinco los imaginamos agrupados de *tres* en tres, descubriremos que definen un icosaedro redondeado (dual del dodecaedro). A tales grupos de tres triangulitos los he llamado icosalitos, de los cuales hay 20, cada uno con una única agrupación de tres colores. Los icosalitos son las unidades elementales e indescomponibles, y equivalen a los cubículos del Cubo o a los bloques elementales de la Pyraminx. Mientras que en el Cubo hay tres tipos de cubículos (aristas, caras y vértices) aquí los icosalitos son de un único tipo. Por esta razón la bola es menos prohibitiva de lo que en principio pudiera temerse. En su estado prístino cada dodecalito es de un solo color. Mèffert se ha valido de seis colores

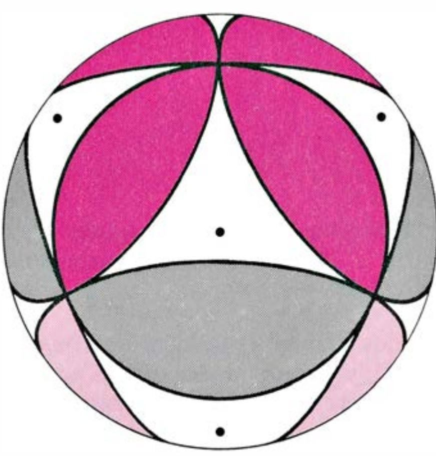


La Parrilla, ideada por Lorente

Otra Parrilla, también de Lorente



El Trébol, de Lorente



La Esfera Florida, de Lorente

(y no doce), empleando un mismo color en los pares de dodecalitos antípodas. Lo que no añade más dificultad al rompecabezas.

Al girar, la I-Bola resulta un tanto sorprendente. Cada grupo de cinco icosalitos concurrentes en un vértice forman lo que llamo un "círculo". Esta agrupación gira como un todo, a razón de 72 grados en cada uno de los dos sentidos cada vez. Cinco de tales giros devolverán al grupo a su posición inicial. Los "círculos" son análogos a las "capas" del Cubo (una cara más su orla). Sin embargo, el "círculo" definido por la frontera de cinco icosalitos no es verdaderamente circular, y si los icosalitos estuvieran sujetos rígidamente, a distancia fija del centro, sería sencillamente imposible hacer girar el grupo de piezas. El mecanismo de Mèffert logra ingeniosamente rodear el escollo, dejando que los icosalitos se alcen ligeramente al pasar sobre "jibas", y el cuerpo del rompecabezas ofrece una flexibilidad perceptible. Por ello, al girar la I-Bola se tiene la deliciosa sensación de estar palpando algo "orgánico".

Las restricciones sobre movimientos son las habituales: todas las permutaciones tienen que ser pares; ello implica que no se pueden trocar entre sí dos icosalitos. Lo más que se puede hacer es permutar cíclicamente tres de ellos (un "ciclo") o permutar simultáneamente dos pares. Como es obvio, la suma de quarks y antiquarks tiene que dar una torsión total entera. Teniendo en cuenta estas restricciones, he calculado que el número total de estados es de 25.563.902.142.421.896.679.424.000, es decir, unos 24×10^{24}, alrededor de 24 millones de trillones, que no alcanza a ser un millón de veces mayor que el número de estados del Cubo de Rubik.

¿Cuál es la dificultad de este rompecabezas? ¿Será aún más complejo que el Cubo? A mi parecer es más fácil, pero sé que puedo no ser justo, porque antes resolví el Cubo. Con terminología de Durham, la IncreíBola es decididamente un rompecabezas "somero", y ello significa que para resolverlo puede bastar un ataque "local". Por mi parte, descubrí pronto que al relajar un poco la minuciosa exactitud conceptual de los operadores que con tanto esfuerzo logré poner a punto en el Cubo, tomándolos no en sentido estricto sino metafórico, parte de la pericia allí adquirida podía trasladarse a la I-Bola. No todo es transferible; inútil es decirlo. Lo que más me gratificó fue comprobar que mis "quarkilladores" y "antiquarkilladores" sí eran directamente exportables. Claro, me hizo falta un rato para determinar en qué podría consistir exactamente tal exportación. ¿Cuál es ahora la esencia de los movimientos? ¿Cuáles de sus aspectos son "provinciales"? ¿Cómo aprender a decidirlo rápidamente? Son éstas cuestiones difíciles, para las que no tengo respuesta.

Gradualmente fui aprendiendo a desenvolverme con la IncreíBola. Me dí cuenta de que una potente categoría de movimientos es la formada por giros de dos "círculos" traslapados, conforme a la pauta de su conmutador ($x\ y\ x'\ y'$). Por consiguiente, estudié primero sobre el papel estos conmutadores, hasta que dí con algunos que cumplían todos los objetivos requeridos [*véase la ilustración de la página anterior*]. Entre ellos se encuentran dobles quarkilladores, dobles transposiciones y tri-ciclos, que forman ya la base de una solución completa. Al hacerlo preparé una notación, la indispensable para mis necesidades, aunque no alcanza a ser una notación completa para la IncreíBola. He aquí lo que a mi juicio sería verdaderamente útil: una notación universal normalizada, satisfactoria tanto matemática como psicológicamente y válida para todos los rompecabezas de tipo cúbico. Se trata, empero, de un proyecto muy ambicioso, puesto que sería necesario prever todas las variantes concebibles sobre este fértil tema, y ello dista de ser empresa trivial.

Es interesante que mis diagramas de círculos traslapados, parcialmente superpuestos, resulten parientes cercanos de otra familia de generalizaciones del Cubo, descubierta por el físico español Gabriel Lorente. Sus rompecabezas, en su mayor parte planares, consisten en redes de círculos rampantes. Lorente ha bautizado a sus rompecabezas "Parrilla" y "Trébol" [*véanse las ilustraciones de esta misma página*]. En cada uno de ellos podemos hacer girar en parte los círculos; seguidamente, algunas de sus partes se entremezclan y redistribuyen. Al generalizar la idea a superficies esféricas, Lorente produjo un elegante rompecabezas, semejante a la IncreíBola, que él llama la Esfera Florida.

Al fijarnos en los rompecabezas de Lorente, en la IncreíBola, e incluso en el Cubo, empezamos a darnos cuenta de que la esencia de todos ellos está, o

así parece, en estas cadenas de órbitas superpuestas. Más aún, podría sostenerse que el carácter tridimensional de estos rompecabezas es irrelevante; su interés reside, a grandes rasgos, en las propiedades de estas órbitas cerradas e intrincadamente encadenadas en un espacio bidimensional, posiblemente curvado como una esfera.

En su forma más pura, el rompecabezas planar de círculos superpuestos fue inventado, según parece, allá por 1890, si bien en tiempos recientes ha sido redescubierto una y otra vez como consecuencia del trabajo en el Cubo. Básicamente, todos estos rompecabezas están formados por dos círculos de cuentas o canicas que se interceptan en varios puntos [*véase la ilustración superior de esta página*]. Podemos desplazar cíclicamente las cuentas de cualquiera de los circuitos; las canicas situadas en las intersecciones quedarán entonces absorbidas por el círculo en movimiento.

Y ya que estamos examinando el carácter bidimensional de estos objetos, vale la pena señalar que la construcción interna de la IncreíBola permite, bastante sorprendentemente, transformarla en lo que llamo rompecabezas 19: una versión bidimensional sobre un espacio curvo del famoso rompecabezas de 15 piezas de Sam Loyd (un cajetín cuadrado de 4 por 4, en donde están encajadas 15 piezas cuadradas y un hueco cuadrado, todos de igual tamaño. El problema consiste en reordenar las 15 piezas numeradas a base de ir cambiando el hueco de lugar). Los primeros que cayeron en la cuenta fueron Halpern y Mèffert. Descubrieron que suprimiendo un solo icosalito (lo que es factible, por ser una de las más bellas características de la IncreíBola que su mecanismo es fácil de montar y desmontar) el hueco dejado por él puede ahora vagar por toda la esfera, exactamente igual que el hueco cuadrado del rompecabeas de Sam Loyd [*véase la ilustración inferior derecha de la página 109*]. Como antes, esta propiedad parece subrayar la naturaleza bidimensional de estos rompecabezas.

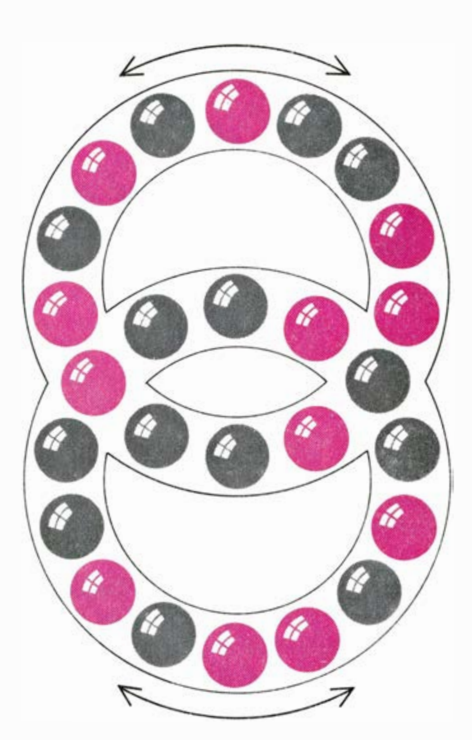
Si se consideran bidimensionales estos rompecabezas es porque en ellos sólo se mueven piezas situadas sobre su superficie; no se producen intercambios entre su interior y su exterior. Como ejemplo extremo, imaginemos que la Tierra sea uno de estos rompecabezas, formado por trillones de círculos de cuentas entrelazados y parcialmente superpuestos. Tal vez con cien millones de giros se pudiera expedir una canica desde La Coruña hasta Nueva York. Tenemos, en esencia, un rompecabezas bidimensional. La pequeñez de los aros con respecto al tamaño terrestre así lo hace obvio. (Sin duda, a nadie se le ocurrirá siquiera la idea de *resolver* semejante problema.)

Examinemos, por el contrario, dos objetos que pronto saldrán al mercado: el cubo de 4 × 4 × 4, fabricado por Ideal, llamado "la revancha de Rubik" (¡no hay que derrochar imaginación!) y la Pyraminx Definitiva de Mèffert, un cubo de 5 × 5 × 5 con vértices y aristas biselados [*véanse ilustraciones de la derecha*]. (Incidentalmente, para los lectores impacientes que dispongan de un ordenador personal existe un programa, el Quadcube, que simula, con gráficos a color, un cubo de 4 × 4 × 4. El programa es compatible con los ordenadores 99/4 y 99/4A de Texas Instruments, y con el ordenador personal de IBM. Lo produce Linear Aesthetic Systems, P.O.Box 23, West Cornwall, Connecticut 06796, Estados Unidos, y cuesta 14,95 dólares por cassette). Estos dos objetos exhiben círculos a escala global mucho mayor. El cubo de 4 × 4 × 4 tiene Círculos Artico y Antártico y Trópicos de Cáncer y Capricornio; el de 5 × 5 × 5 tiene además un Ecuador.

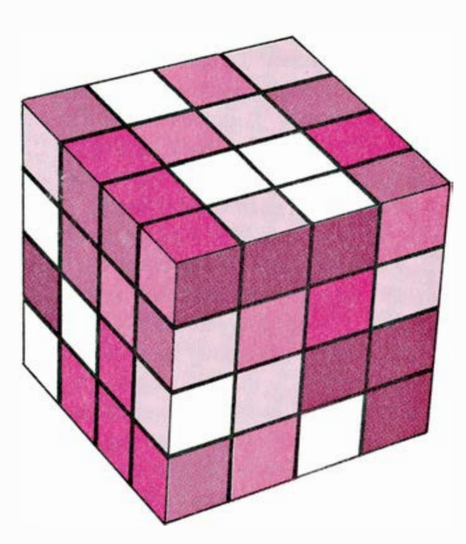
En el cubo de 3 × 3 × 3 podemos despreciar el Ecuador, porque los giros ecuatoriales equivalen a giros de los casquetes polares, algo así como hacer girar las rebanadas de pan de un bocadillo sin hacer girar el jamón. La notación elegida por Singmaster para el cubo 3 × 3 × 3 refleja su propensión a considerar estáticos los centros de las caras. Según él, las rebanadas se mueven, y el jamón ha de permanecer en su sitio. En principio, espléndido. En la práctica, la gente no mantiene el bocadillo –perdón, el Cubo– constantemente en la misma posición. Además, al aumentar el orden, este planteamiento resulta insuficiente. Imaginemos un sandwich de varias capas, con tres rebanadas de pan y dos clases de carne intermedias. Sencillamente: es necesario ampliar el horizonte notacional.

En su libro *Winning Ways* (Academic Press, 1981), John H. Conway, Elwyn R. Berlekamp y Richard K. Guy han propuesto una elegante nomenclatura para los seis movimientos ecuatoriales del cubo 3 × 3 × 3. Emplean letras griegas con muy bien pensadas justificaciones mnemónicas [*véase la ilustración superior de la página siguiente*]. Con algunas modificaciones, estas letras podrían adaptarse a las rebanadas de cubos de orden superior.

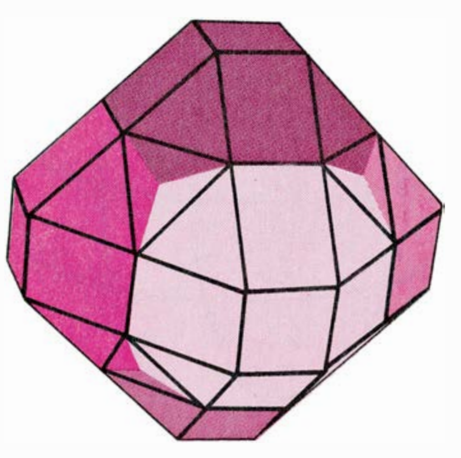
Los movimientos de las rodajas de estos cubos de categoría superior vienen a ser como gigantescos círculos de



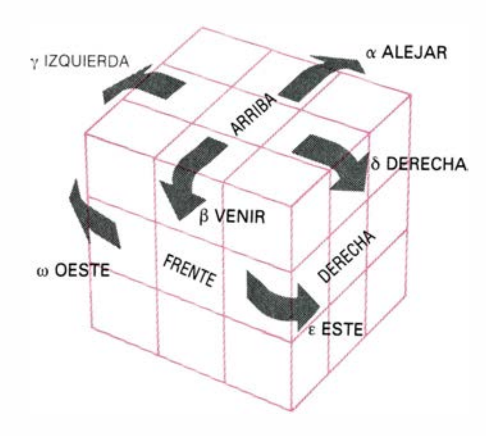
Rompecabezas de círculos secantes



El Cubo de 4 × 4 × 4, desordenado



La Pyraminx Definitiva



Notación de los movimientos ecuatoriales del Cubo 3 × 3 × 3

canicas que contorneasen el Ecuador, o el Trópico de Capricornio. Sus radios son del mismo orden de magnitud que el radio del objeto tridimensional subyacente. La topología de los enlaces y ligaduras entre círculos es ahora mucho más complicada que cuando los círculos eran pequeños, y las conexiones entre ellos son de carácter muy local. Para describir económicamente el sistema de ligaduras tendríamos que referirnos a la forma en que los círculos están inmersos en el espacio tridimensional. En este sentido, para los rompecabezas de orden superior sí puede decirse que son intrínsecamente tridimensionales.

Conforme se le dan vueltas al Cubo, de las salpicaduras se van creando, o así parece, otros innumerables rompecabezas. La idea que engendró al Cubo ha sido muy fecunda. He recibido un precioso artículo de H. J. Kamack y T. R. Keane, quienes han simulado un cubo de 3 × 3 × 3 × 3 en el espacio tetradimensional con ayuda de ordenador, al que han bautizado Rubik's Tesseract. El número de sus distintos estados es casi un gúgol (10^{100}). (Si Ideal Toy se decidiera a comercializarlo podrían decir que tiene "más de tres trillones de combinaciones".) Kamack y Keana han hecho muchos incitantes descubrimientos, pero no puedo dar aquí noticia de ellos. Los lectores interesados pueden dirigirse a Kamack, escribiéndole al 610 Lehigh Road (Apartment 10), Newark, Delaware 19711, Estados Unidos.



Penoso estado de nuestro mundo

He recibido también un fascinante artículo de György Marx y Eva Gajzágó, ambos físicos de la Universidad Roland Eötvös de Budapest. Dan en él una definición de "entropía" para el Cubo, y describen algunos resultados estadísticos determinados por un estudiante de secundaria llamado Victor Zámbó. Son cuestiones de las que me gustaría hablar en el futuro.

Para terminar, me gustaría comentar la increíble popularidad alcanzada por el Cubo. En la lista de títulos más vendidos que publicaba el *New York Times* del 15 de noviembre último figuraban tres libritos sobre el Cubo, ¿y en qué puestos? ¡En los números 1, 2 y 5! La gente suele preguntarse, "¿Por qué es el Cubo tan popular? ¿Será duradero, o sólo el furor de un día?" A mi juicio, el Cubo perdurará. Me parece que tiene una especie de atractivo "primordial", básico e instintivo. Conceptualmente, su "gancho" le viene de hallarse alojado en un receptáculo de la mente humana que conecta con muchas nociones generales concernientes al mundo. Cuanto sigue es un intento de caracterizar tal cualidad.

Para empezar, el Cubo es pequeño y rico en colorido. La mano lo ciñe fácilmente; es grato al tacto. La torsión es un movimiento fundamental e intrigante, que la mano ejecuta con naturalidad. El objeto propiamente dicho ofrece simetría general, y por ello podemos darle vueltas a su todo sin sentirlo cambiar de forma. (A diferencia de la mayor parte de los rompecabezas, que tienen, cuando más, un eje de simetría.) Es curioso, pero son pocos los rompecabezas que proporcionan a la mente y los dedos un genuino ejercicio tridimensional.

Incluso al desordenar el objeto éste permanece siendo un todo (a diferencia de muchos otros pasatiempos que, como Humpty-Dumpty, se desmoronan en cascadas de piezas que luego quedan esparcidas por los suelos). Que el objeto logre permanecer conexo a pesar de las muchas formas independientes de torsión que admite causa auténtica admiración al principio, y sigue resultando misterioso incluso tras haber penetrado en su intimidad.

El objeto es encarnación miniaturizada de esa sutil mezcla de orden y caos que es nuestro mundo. Casi nunca puede uno predecir qué repercusiones tendrán las acciones más simples; sencillamente, hay demasiados efectos secundarios. Unos cuantos actos pequeños y sencillos pueden tener vastas consecuencias mutuamente entrelazadas, y resultar impracticables. Es fácil que el temor nos paralice, incapacitándonos para movimiento alguno, sabiendo que sin sentirlo podemos llegar a perdernos total, desesperada e irremisiblemente.

El Cubo ofrece enorme riqueza de configuraciones, algunas accesibles, otras no. Algunas se engendran fácilmente, pero no sabemos cómo. Otras son difíciles de generar, pero comprendemos claramente cómo se producen.

Hay muchas rutas conducentes a un mismo estado, y, de ellas, la más breve es casi siempre imposible de conocer. La solución de una dificultad raramente consiste en retornar por el camino que nos llevó a ella, sino en dar con una ruta de escape alternativa y totalmente distinta. Se siente uno como atrapado en una cueva, sin luz, incapaz de percibir el espacio completo, sólo la vecindad más inmediata, preguntándose si será humanamente posible alcanzar la visión general.

Es el Cubo rica fuente de metáforas. Nos provee de analogías con la física subatómica (de los quarks, pongamos por caso) con la biología (una sucesión de movimientos sería "un genotipo", la configuración codificada por ella, el "fenotipo"), con la resolución de problemas de la vida ordinaria (descomposición del problema en partes, y resolución individual de cada etapa), con la entropía, la exploración e incluso la teología ("algoritmo de Dios").

La comprensión del Cubo puede lograrse desde distintos enfoques. Hay, en particular, un fuerte contraste entre los enfoques "algebraico" y "geométrico". En el enfoque algebraico, o algorítmico, a partir de secuencias breves de operaciones se construyen secuencias largas, con lo que al cabo de un momento no se tiene idea de por qué estamos ejecutando los giros indicados; sencillamente, confiamos en que la secuencia completa funcione como un todo. Es un método eficiente, pero arriesgado. En el enfoque geométrico, de sentido común, ojo y mente colaboran para ir decidiendo un giro tras otro, cada uno con clara razón de ser, formando parte de una ruta minuciosamente señalizada. Es método de fiar, pero de baja eficiencia. Ambos enfoques pueden servir válidamente como símiles de los estilos en la resolución de problemas.

El universo del Cubo tiene una curiosa población. Aparte de las variedades de cubículos y modos de giro, existen también efímeras entidades que podríamos llamar "revirado" y "retorcimiento", a las que uno va literalmente haciendo recorrer el Cubo, lo mismo que vamos trasladando físicamente sus cubitos materiales y tangibles. Análogamente, la palabra "aquí" puede designar un "lugar" que va de acá para allá durante una sucesión de giros. Los sistemas de referencia, entrelazados y subyacentes entre sí, y entre los que tenemos que ir saltando para tratar de recomponer el Cubo sirven vívidamente de ejemplos de la estratificación de niveles con que concebimos el espacio, y más aún, de la propia estratificación de nuestro pensamiento.

El Cubo ofrece también otros encantos, de orden no tan intelectual; entre ellos, la magia del movimiento más rápido que el ojo, la emoción de la competición de alta velocidad, el conocimiento de los distintos grados de perfección que con él podemos alcanzar, los gozos del intercambio de información e ideas. E incluso tras largo rato de jugar con él, uno se pregunta cómo un objeto tan pequeño e inocente puede ocultar tan vasto universo de potencialidades.

Finalmente, fijémonos en la metáfora que el Cubo ofrece acerca del estado del mundo. (Es metáfora muy explotada en el humorismo político.) El globo está revuelto [*véase la figura inferior de la página 114*] y los líderes de los diversos países desean "arreglarlo". No están dispuestos, sin embargo, a sacrificar una pizca del poco de orden que consiguen. Se aferran a logros caducos e inoperantes por su excesivo temor a dejarlos escapar, perdiendo temporalmente la ordenación que ya tienen y a la que es forzoso renunciar para buscar mayor ordenación y armonía. ¿Será que carecen de visión global, del enfoque superior, capaz de reconocer que la disposición a ciertos sacrificios a la corta puede comportar beneficios mucho mayores a la larga?

Yo tengo confianza en que el Cubo, y los "cubos" en general, llegarán a florecer. Espero que durante mucho tiempo sigan apareciendo nuevas variedades, que enriquezcan nuestras vidas de muchas y nuevas formas. Resulta muy grato que un juguete tan exigente y estimulante haya alcanzado tan gran éxito mundial. Me han dicho que ya es muy popular en China. Tal vez consiga traspasar las fronteras de la Unión Soviética, que, por lo que sé, debe ser el último bastión del mundo Cubo-libre.

Taller y laboratorio

De un caramelo que se ilumina en la boca y de los curiosos rebotes de la luz en una lente

Jearl Walker

Si el lector se coloca frente a un espejo dentro de una habitación a oscuras durante el tiempo suficiente para que su vista se adapte a la oscuridad y, seguidamente, comienza a mascar un trozo de caramelo duro, llegará a observar en la imagen reflejada de su boca una luz tenue parecida al fulgor de la esfera de un reloj luminiscente. Ese fulgor es una muestra de la triboluminiscencia (del griego *tribein*, frotar), fenómeno que puede ponerse de manifiesto con la mayoría de los caramelos duros elaborados con azúcar granular. De él es responsable el azúcar; o, más exactamente, las moléculas de nitrógeno gaseoso que lo rodean.

La primera vez que reparé en el fenómeno fue ante un espejo en una habitación a oscuras, mientras mascaba un caramelo de gaulteria. Al triturarlo con los dientes pude observar en la imagen que, en el interior de mi boca, resplandecía como una luz azulina. El hecho se conoce, por lo menos, desde el siglo XVII, cuando la Accadèmia del Cimento de Florencia publicó una descripción del mismo que, en lenguaje actual, dice: "De los Cuerpos generadores de Luz... Además de los Pedernales existen otros Albergadores de Luz; pues, al entrechocarlos o partirlos en la Oscuridad, relucen. Tales son el Azúcar Blanco, el Azúcar de Pilón y la Sal Gema; todos los cuales, cuando se machacan en un Mortero, despiden una Luz tan intensa que, con ella, pueden distinguirse perfectamente los costados del Mortero y la forma de la Mano de Almirez." Casi al mismo tiempo Robert Boyle advertía que "el azúcar duro, al ser raspado vivamente con un cuchillo, despide una luz chispeante". Desde entonces, son numerosos los investigadores que han indagado en el fenómeno, pero la naturaleza de la triboluminiscencia no ha sido conocida hasta nuestro siglo. Y, aún hoy, se reserva puntos oscuros.

Hasta hace muy poco el estudio de la triboluminiscencia se centraba en averiguar qué sustancias lo presentan, examinar los espectros luminosos y analizar de qué modo correlacionar el fulgor con las propiedades físicas de las sustancias. Gran parte de las investigaciones modernas se deben a Jeffrey I. Zink y sus colaboradores, de la Universidad de California en Los Angeles. Estos, y otros, han averiguado cuál es el origen de la luz en el azúcar y varias sustancias más. Sin embargo, el esclarecimiento de la relación entre la luminiscencia y la compresión ejercida sobre la sustancia sigue siendo un tema escurridizo.

Triboluminiscencia la exhiben un buen número de sustancias, pero no de la misma forma que el azúcar común (sucrosa) en todos los casos. Zink y otros han encontrado que azúcares como la D-glucosa, lactosa, maltosa y L-ramnosa la poseen, mientras que otros, como la fructosa, celobiosa, fucosa, galactosa y manosa, no.

A principios de siglo se demostró que el espectro de la luminiscencia producida por la sucrosa era igual al del nitrógeno gaseoso. Revelose, pues, que no se originaba en las moléculas del azúcar, pues éstas habrían emitido un espectro completamente diferente. El espectrógrafo muestra que la triboluminiscencia del azúcar posee varios picos brillantes en las proximidades de la zona ultravioleta del espectro. En los 330 y 360 nanometros aparecen emisiones intensas y otras más débiles en los 315 y 380 nanometros. La luminiscencia visible es más apagada y yace entre los 380 y 415 nanometros (zonas del violeta y azul del espectro). Si el ojo humano fuese sensible al ultravioleta, el fulgor sería más brillante ya que en esa hipótesis se visualizarían los picos de emisión intensos.

Al machacar cristales de azúcar, por alguna razón se excita el nitrógeno molecular asociado al azúcar. El estudio del espectro revela que el nitrógeno se excita desde su nivel energético molecular más bajo hasta el nivel $^3\pi_u$ e, inmediatamente, cae al estado $^3\pi_g$, emitiendo luz compuesta de múltiples longitudes de onda. Estas se distribuyen según el patrón que caracteriza, de modo exclusivo, al nitrógeno molecular.

Otro método para evidenciar el papel del nitrógeno consiste en eliminar del azúcar la mayor cantidad posible de aquél. Así, cuando se aplasta azúcar en atmósfera muy enrarecida, la luminiscencia es menos intensa, pero persiste todavía ya que, incluso en el vacío, quedan algunas moléculas de nitrógeno adsorbidas en, o absorbidas por, el azúcar. Cuando el nitrógeno se elimina del todo, tal como Zink ha podido conseguir, desaparece la luminiscencia.

Análogamente se revela la importancia del gas ambiental si sustituimos el nitrógeno por otros gases. Al introducir azúcar en una atmósfera de neon a baja presión, la luminiscencia por aplastamiento se torna roja. ¿Por qué? La compresión excita, de un modo u otro, al neon y, luego, éste se desexcita, más o menos como en un anuncio luminoso. El neon emite intensamente en la zona roja del espectro.

Determinados caramelos duros despliegan colores adicionales. Zink y sus colaboradores han hallado, por ejemplo, que los caramelos de gaulteria despiden una tonalidad más azul verdosa que los de otros sabores, los cuales presentan la emisión azul pálido propia de la sucrosa. Ello obedece a que el metil-salicilato (aceite de gaulteria) con que se sazona el caramelo absorbe la radiación ultravioleta del nitrógeno. Cuando el metil-salicilato se desexcita, emite en un amplio espectro comprendido entre los 400 y los 540 nanometros, aproximadamente. El masticador que lo compruebe verá entonces la combinación de los espectros del nitrógeno molecular y del metil-salicilato.

Me he dedicado a producir el efecto de la triboluminiscencia moliendo azúcar de mesa y varios productos azucarados más en un cubilete, previa adaptación de mi vista a la oscuridad. No parece que importe mucho el útil con que se muela; así, me han dado buen resultado tanto el extremo romo de un cuchillo de mesa como una espiga de madera. Ni importa que el cubilete sea de vidrio, plástico o metal. (Cuidado con el vidrio; trabajando en oscuridad, podría romperse el recipiente si se aplica la presión de modo incorrecto.)

La abundancia de la luz emitida decrecía a medida que proseguía moliendo. El azúcar ligeramente contaminado por la humedad ambiental despedía una luminiscencia más tenue. El disuelto en agua no daba centelleo alguno, a no ser que percutiera algún cristal no disuelto con el útil de moler.

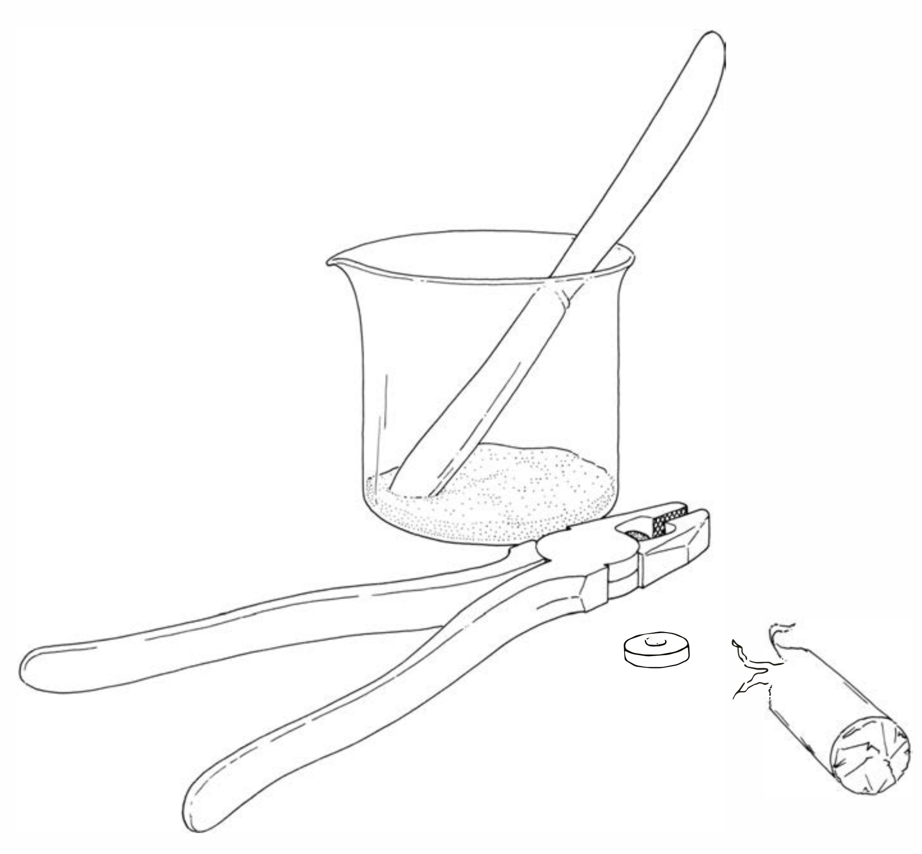
Luminiscencia de un caramelo de gaulteria aplastado con unos alicates

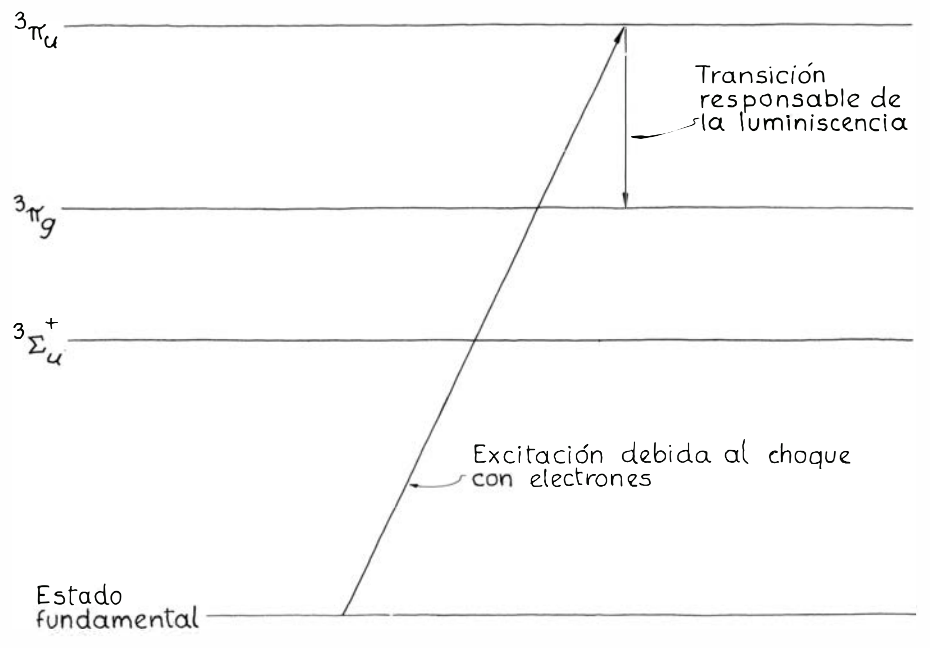
Solución de azúcar que actúa de lente gradiental en el experimento de William M. Strouse

Numerosos detergentes para ropa líquidos y en polvo contienen un ingrediente que absorbe el ultravioleta y emite en el espectro visible. (Con la intención de que el producto "lave más blanco".) Añadí uno de esos productos al azúcar, pero no observé diferencia alguna.

También me ocupé de la luminiscencia producida por caramelos de diversos sabores. Como protección para mi dentadura, seguí la sugerencia de Zink y trituré primero los caramelos con unos alicates. Los de gaulteria creaban la luz más fácilmente visible, sin duda a causa de la emisión del metil-salicilato.

Para estudiar ese espectro pegué con cinta adhesiva una red de difracción a mis gafas. Mirando directamente al fulgor recibía el pico central de la figura generada por la red. Pero en este pico se solapan todos los colores procedentes de un manantial luminoso, por lo que no es posible la determinación del espectro. El primer pico lateral procedente de una red de difracción descompone en sus colores la luz que surge de un manantial, pero no pude entonces distinguirlo porque la luz que despedía el caramelo era muy débil; ni siquiera podría apreciarse un pico lateral más luminoso, ya que, al ser tan débil la luz, la retina tendría quizá dificultades para distinguir los colores.


Material necesario para el estudio de la triboluminiscencia



Niveles de excitación y desexcitación del nitrógeno responsables de la luminiscencia

Una red de difracción como la que les he citado puede adquirirse en Edmund Scientific Company (101 East Gloucester Pike, Barrington, N.J. 08007). Quienes hayan construido alguno de los analizadores espectrales descritos en mis artículos anteriores, lograrán separar el espectro de un trozo de caramelo duro. En el artículo de Rebecca Angelos, Zink y Gordon E. Hardy que se cita en la bibliografía [*página 128*] se expone un procedimiento para realizar la experiencia.

Triboluminiscencia la he encontrado en las pastillas de sacarina, cristales de ácido tartárico y sal gema. La sal de mesa y los caramelos blandos no la dan. La razón tal vez sea la siguiente: aunque los caramelos blandos contengan posiblemente azúcar, los cristales del mismo serán demasiado pequeños o demasiado difíciles de someter a esfuerzos mecánicos (dada la plasticidad del caramelo) para que aparezca ese efecto.

Los caramelos duros desprenden a veces luminiscencia en virtud de una conmoción repentina motivada por una causa distinta de la comprensión. Así, cuando un caramelo se deja caer en nitrógeno líquido, la reducción súbita de temperatura crea una tensión mecánica suficiente para generar un fugaz destello. (Si alguien repite el ensayo, que ande alerta con el nitrógeno; puede resultar muy peligroso si se maneja incorrectamente.)

Muchas otras sustancias ofrecen una triboluminiscencia originaria de la propia sustancia y no del gas ambiente. Nos hallamos en ese caso cuando el espectro de la triboluminiscencia es igual al que emite la sustancia al ser calentada o iluminada. Volvemos a encontrar aquí que la relación entre la compresión y la excitación de las moléculas emisoras no se conoce bien. Podría ser consecuencia de reacciones químicas. Y pudiera ser térmica también, origi-

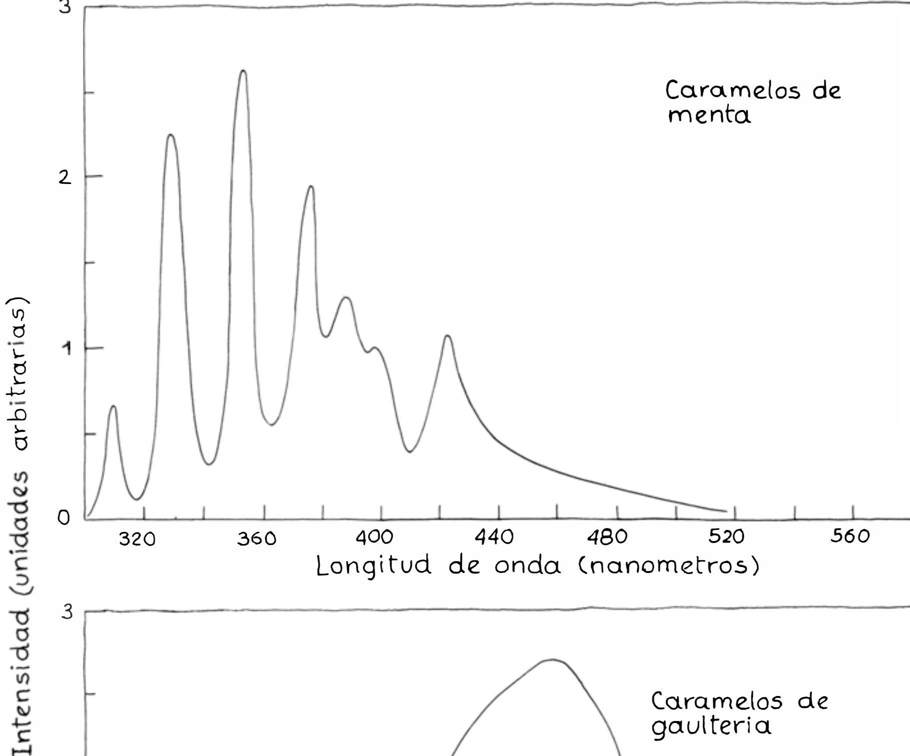
nándose porque la compresión deforme las moléculas, altere las interacciones entre moléculas contiguas o desplace las dislocaciones existentes en los cristales de la sustancia.

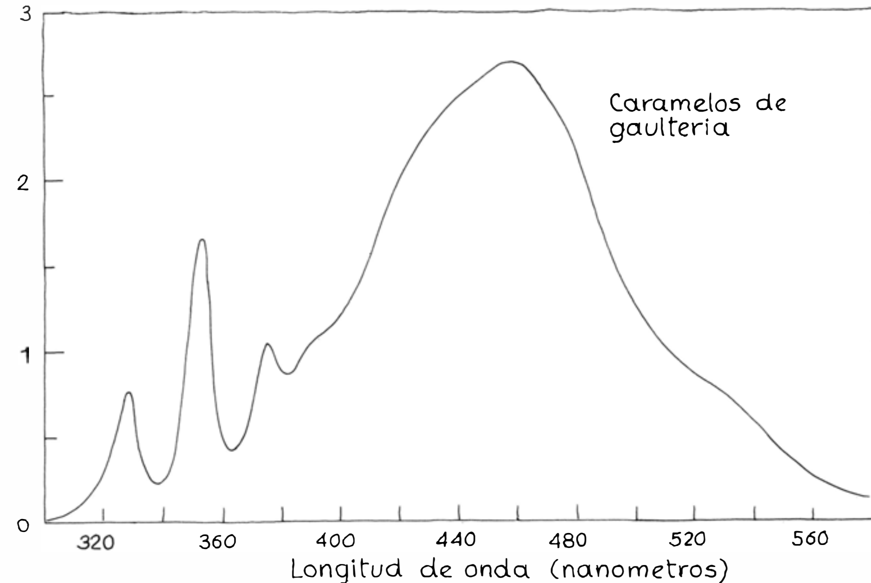
La luminiscencia de los caramelos duros parece deberse a descargas eléctricas y no a mecanismos químicos ni térmicos. Las descargas eléctricas resultan de la separación de cargas que provoca la tensión mecánica aplicada a la sustancia. Efectivamente, aunque una sustancia no se halle fuertemente cargada, el esfuerzo mecánico que recibe libera iones y deja electrones semilibres. Entonces, la luminiscencia sería consecuencia de la recombinación de iones positivos y negativos. También podría ser que se origine al combinarse los electrones semilibres con lugares positivos de la sustancia, desprendiendo luz. Una tercera posibilidad es que los electrones sean acelerados por los campos eléctricos existentes entre los lugares con carga y choquen con las moléculas y las exciten, por lo que emitirán luz al caer a niveles energéticos inferiores.

En el caso del azúcar, la luminiscencia parece achacable a un efecto eléctrico cuyo resultado es la excitación molecular por choque. Cuando se comprime azúcar, en distintos puntos de su interior se liberan electrones. Estos cruzan de un salto cada cristal de la sustancia, o de un cristal a otro, buscando lugares cargados positivamente. En este proceso chocan con el nitrógeno molecular adsorbido en o absorbido por el azúcar y, además, con moléculas de nitrógeno gaseoso. Cualquiera que sea el caso, los electrones ceden parte de su energía cinética al nitrógeno molecular, elevando a éste a un estado excitado.

En el azúcar la creación de superficies cargadas parece deberse a la fractura de los cristales de azúcar. Toda fractura deja superficies cargadas que, casi en el acto, son cubiertas por el nitrógeno molecular del aire. Y queda allanado el camino para que los electrones en movimiento arranquen luz del nitrógeno.

Zink y sus colaboradores han investigado la rapidez a la que se propagan las grietas en los cristales de azúcar, encontrando que una grieta simple y su luminiscencia asociada deben persistir alrededor de un microsegundo. Al masticar un caramelo de gaulteria, o al triturarlo con unos alicates se produce una luminosidad más duradera debido a la progresión de las grietas. La luminosidad depende del tamaño de los cristales. La luminiscencia disminuye





Espectrogramas de triboluminiscencia obtenidos por Jeffrey I. Zink y sus colaboradores

conforme se sigue triturando un trozo de caramelo duro, porque los cristales se van haciendo más pequeños.

Otro material que presenta triboluminiscencia: las cintas cubiertas de adhesivo por un lado. E. Newton Harvey, de la Universidad de Princeton, mostró en 1939 que hay varios tipos de cinta que despiden luz en la oscuridad cuando se estira rápidamente un trozo del rollo o se pasa por encima de determinadas superficies, vidrio por ejemplo. El fenómeno se observa también si se pegan dos capas de cinta y luego se separan. Ocasionalmente aparece al pasar un dedo a lo largo de la cara no recubierta de un trozo de cinta adhesiva.

Como en el azúcar, la luminiscencia se debe a una descarga en el seno del nitrógeno gaseoso ambiente. Su espectro coincide con el del nitrógeno molecular. Si la cinta se desenrolla en una atmósfera de neon a baja presión, el fulgor será rojo y no azul.

Parece que la manipulación de la cinta adhesiva engendra pequeñas superficies cargadas. Cuando éstas se hallan separadas, los electrones de las superficies negativas saltan en el seno del gas para alcanzar las negativas, excitando el nitrógeno ambiente por el camino. Algo similar se observa al partir mica y, excepcionalmente, en una banda de caucho estirada cuando se libera de golpe la tensión.

Harvey describió cómo demostrar que, al arrancar un trozo de cinta adhesiva de una base, se produce una separación de cargas. Para ello pegó un trozo sobre una lámina de vidrio; luego la

arrancó. Al sostener la cinta junto al vidrio, aquélla se veía atraída por las zonas cargadas que se habían generado en ambas superficies. Mediante bolitas de médula cargadas, comprobó el signo de la carga de cada superficie. Una bolita cargada positivamente era atraída por la cinta, indicando que ésta se había hecho positiva. Análogamente, una bolita cargada negativamente era atraída por el vidrio. Esto no significa necesariamente que la cinta sea negativa en su totalidad y el vidrio sea positivo, sino que en cada superficie hay bolsas de cargas localizadas en las que predominan las cargas de un signo.

Los dispositivos ópticos tradicionales, como lentes y prismas, trabajan refractando la luz, o sea, desviándola cuando atraviesa sus superficies. La luz se refracta cuando atraviesa la frontera entre dos sustancias en las que se propaga a velocidades diferentes. Una manera para medir hasta qué punto es refringente una sustancia consiste en asignarle un índice de refracción, de tal modo que cuanto más despacio se propague la luz en su seno, mayor sea su índice de refracción. Por ejemplo, el índice de refracción del vidrio oscila en torno a 1,5, mientras que el del aire es ligeramente mayor que 1.

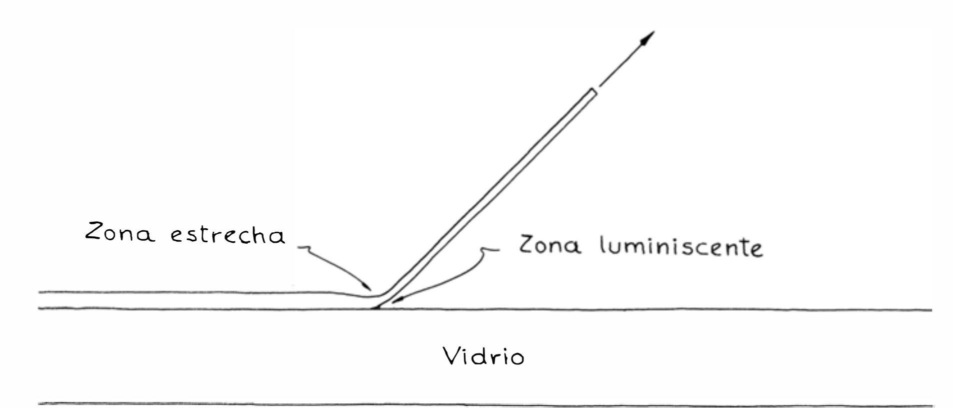
La luz se refracta al entrar y al salir de una lente de vidrio, pero no dentro de ésta, pues el índice de refracción no varía allí. Si las caras del vidrio fueran planas y paralelas, la refracción de la luz en la superficie de entrada se contrarrestaría con la refracción en la superficie de salida. Sin embargo, con el propósito de alterar la dirección de propagación de la luz, se dota a las lentes de superficies curvas; así, las lentes convexas cambian la dirección de los rayos luminosos paralelos, haciendo que éstos converjan en el punto llamado foco, tras atravesar la lente.

Las lentes con gradientales presentan la originalidad de que su índice de refracción no es constante. Hay lentes de este tipo cuyo índice varía perpendicularmente al eje óptico (o eje paralelo a la dirección de la luz). En el centro de la lente el índice de refracción puede ser mayor que en las zonas periféricas. Entonces, cuando la luz la atraviesa, los rayos próximos al eje óptico se desviarán muy poco, pero los más alejados virarán acusadamente hacia el mismo. Y así, aun cuando las superficies de la lente sean planas, podrá enfocarse la luz cual si se tratara de una lente convexa ordinaria. Si tales superficies fuesen curvas, la luz se refractaría todavía más.

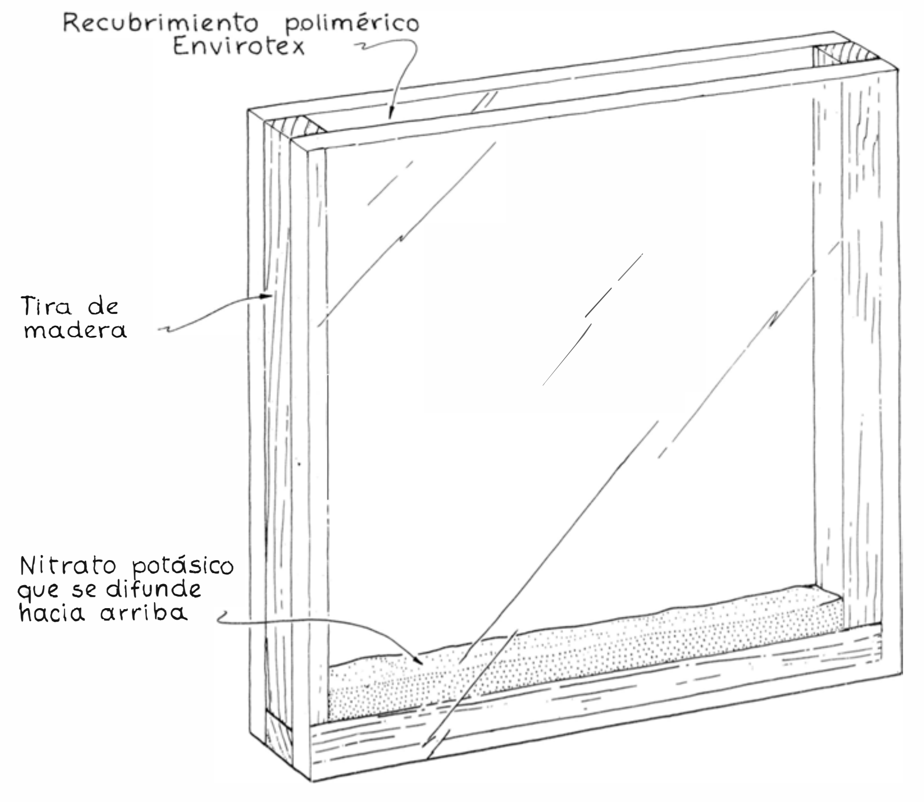
En un segundo tipo de lente con gradiente de índice de refracción, dicho gradiente está orientado paralelamente al eje óptico, y no perpendicularmente a él. O sea, el índice de refracción en la superficie anterior es distinto del que se registra en la posterior, variando gradualmente entre una y otra. En un tercer tipo el gradiente es esférico; por ejemplo, puede que el índice de refracción decrezca hacia fuera, a partir del centro, con simetría radial. Aunque ignoro si han llegado a construirse lentes así, James Clerk Maxwell estudió su óptica hace más de un siglo; un prototipo de ellas es el cristalino del ojo.

Se han encontrado varias aplicaciones sencillas de las lentes con gradientales pero no se han generalizado dada la dificultad de su construcción. Jurgen R. Meyer-Arendt, de la Universidad Pacific, y su discípulo Mark R. Zilm me remitieron recientemente la descripción de un procedimiento fácil para construir una lente gradiental sencilla, con el que han conseguido una que se comporta como una lente cilíndrica cóncava.

Esencialmente, esa lente se compone de epoxia transparente en la que se difunde nitrato potásico al objeto de crear el gradiente de índice de refracción. Dichas sustancias se encierran en un receptáculo formado por dos trozos de vidrio cuadrados, de 50 milímetros de lado y 2 milímetros de grueso cada uno; se mantienen separados y paralelos mediante trocitos de madera delgados (palillos para fósforos) de 50 milímetros de largo y 3 milímetros cuadra-



Lugar donde se genera la luminiscencia al arrancar cinta adhesiva de vidrio



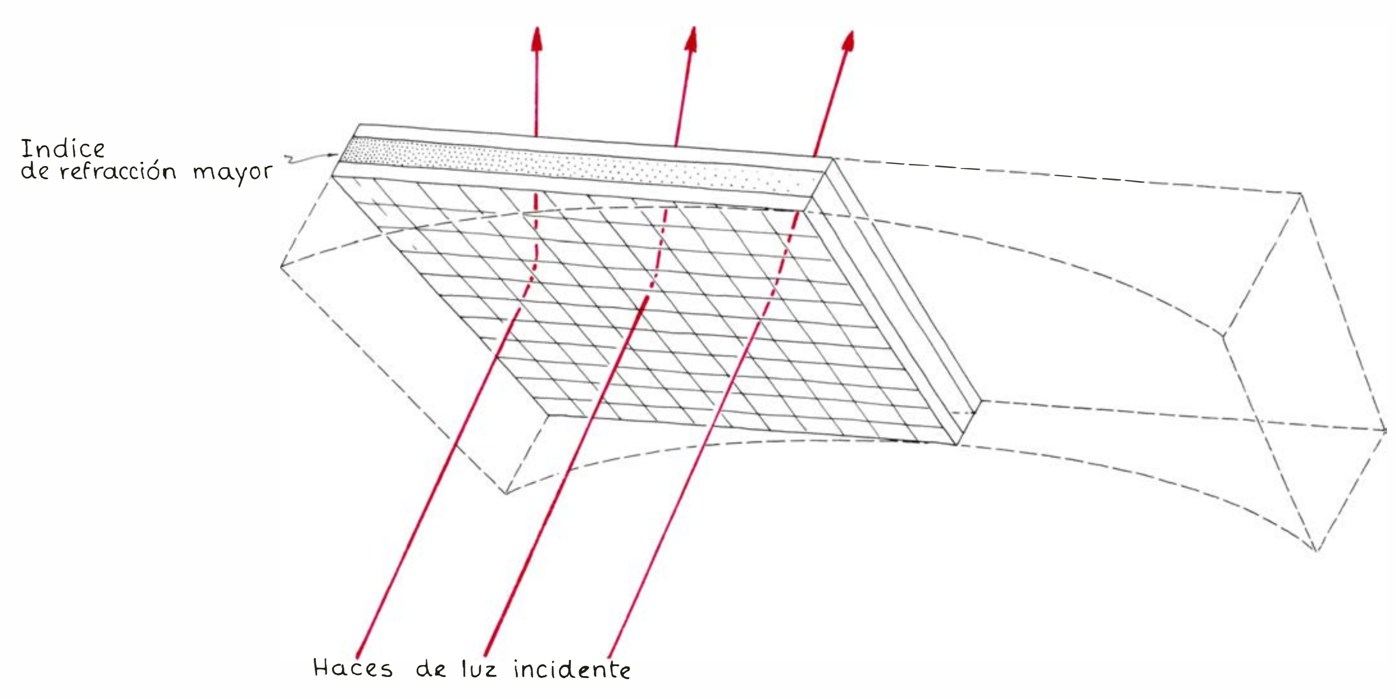
Lente gradiental construida por Jurgen R. Meyer-Arendt y Mark R. Zilm

Indice de refracción mayor
Base original
Desviación
Haz de láser
Base original; índice de refracción mayor
Pantalla
Lente gradiental
Láser
Indice de refracción mayor

Refracción de un haz láser en una lente gradiental

dos de sección. Emplearon en un principio madera de balsa, pero Meyer-Arendt y Zilm descubrieron luego que resultaba mejor la madera dura de los fósforos. Sobre tres cantos de cada trozo de vidrio se encolan los maderitos mediante pegamento sellador de silicio. (Un pegamento de otra naturaleza quizá no diese el mismo resultado, al difundirse en la epoxia.)

El cuarto lado del receptáculo se deja abierto, para verter por él la epoxia y el nitrato potásico (KNO_3). Este, que se conoce también como salitre, puede comprarse en cualquier droguería. La epoxia preferida por Meyer-Arendt y Zilm es recubrimiento polimétrico Envirotex, que se vende en las tiendas donde se expende material para confeccionar adornos de papel recortado. En el fondo del recipiente se dispone una capa uniforme de nitrato de unos dos milímetros, y luego se introduce la epoxia. Meyer-Arendt han mo-



Semejanza entre una lente gradiental y otra cóncava

dificado el procedimiento habitual de mezclar la epoxia, al objeto de que tarde más en endurecerse y el nitrato disponga de tiempo para difundirse. Corrientemente, la relación entre endurecedor y epoxia es de 1:1; ellos han tomado una de 1:4. Al pretender rebajar más el contenido de endurecedor, hallaron que se formaban burbujas de aire en la mezcla al endurecerse ésta.

Tras dejar reposar el receptáculo durante tres semanas, el nitrato potásico se había difundido hacia arriba dentro de la epoxia y toda la mezcla se había solidificado. El resultado fue una lente cuyo índice de refracción varía de abajo arriba, siendo mayor el índice de refracción en el fondo, donde también es mayor la concentración de nitrato, y menor en la parte superior, donde no penetró el nitrato.

Meyer-Arendt y Zilm han ideado un experimento sencillo para poner de manifiesto las características refringentes de su ingenio. Primero, grabaron químicamente una retícula sobre una de las piezas de vidrio laterales del dispositivo. Esa retícula está formada de rectángulos de tres milímetros de alto por cuatro de ancho. Luego colocaron la lente de canto, de modo que quedara vertical lo que antes era base.

La operación siguiente consistió en dirigir un haz láser sobre una zona de la retícula. La luz, tras atravesar la lente, iba a parar a una pantalla situada a diez metros. Comprobando primero el punto de la pantalla donde incidía el haz con la lente retirada y, luego, el punto donde incidía con la lente en posición, Meyer-Arendt y Zilm midieron la desviación provocada por la lente. Repitieron esas mediciones a lo largo de una serie de zonas de la retícula alineadas y con índices de refracción decrecientes, desde el correpondiente a la porción de lente con el índice máximo hasta la de índice mínimo.

En la ilustración superior de la página anterior puede verse la refracción de un haz láser al atravesar la lente gradiental por la proximidad de lo que en un comienzo fue el fondo. Puesto que el haz se propaga en el seno de un gradiente de índice de refracción, la parte del mismo que atraviesa la superficie que presenta un índice mayor se rezaga respecto a la que atraviesa la superficie de índice menor y, en consecuencia, se desvía hacia el lado que antes fue la base.

Tal como se aprecia en la ilustración inferior, las propiedades de esta lente gradiental recuerdan a las de una media lente cóncava. En las zonas donde el gradiente de concentración de nitrato potásico es muy acusado, la luz se desvía mucho, tal como ocurriría en las zonas más externas de una lente cóncava. En las zonas donde el gradiente de concentración de nitrato es menor, la luz se desvía casi como lo haría en el centro de la lente cóncava. Si pudiera controlarse exactamente la difusión del nitrato potásico en la epoxia, se construirían lentes gradientales equiparables a cualquiera de las ordinarias.

Si alguien fabrica una lente con gradiente de índice, siguiendo el método de Meyer-Arendt y Zilm, quizás encuentre interesante explorar sus características cromáticas. En tal caso, a no ser que tenga acceso a luz láser de colores variables, habrá de trabajar con la luz parcialmente coherente procedente de un estenope. De manantial podría servirle cualquier luz intensa filtrada para estrechar el espectro de colores del haz. Necesitará un ocultador o un diafragma para confinar la luz en sólo una pequeña superficie de la lente.

En la bibliografía se cita un experimento curioso de refracción gradiental que realizó William M. Strouse. Consistía en dirigir un haz luminoso a través de una pareja de lentes (al objeto de estrecharlo) y, luego, al interior de un recipiente lleno de una solución de azúcar en agua. El recipiente tenía dos centímetros de ancho por ocho de alto y veinte de largo. Strouse introdujo en él tres centímetros de agua, a la que añadió tres terrones de azúcar, que dejó disolverse tranquilamente. El azúcar, al difundirse por el agua, creó un gradiente de concentración de modo que ésta era mayor en el fondo y menor en la superficie libre del líquido, en virtud de lo cual el índice de refracción variaba de mayor a menor en el mismo sentido.

Cuando dirigía el haz sobre la solución cerca del fondo del recipiente, se curvaba hacia abajo y se reflejaba en el fondo. Pero el haz reflejado seguía curvándose y, si el recipiente hubiese sido largo, lo habría recorrido rebotando repetidamente en el fondo.

Visto desde arriba, el haz conserva la anchura al atravesar el recipiente. Visto de costado, se estrecha periódicamente. Hay que achacar esa focalización parcial al gradiente de índice de refracción. El gradiente altera la dirección de los rayos en proporciones distintas, de suerte que el haz se estrecha en ciertas zonas de la solución y se ensancha en otras. Strouse señala que, si queremos realizar el fenómeno, retiremos las lentes destinadas a estrechar el haz original, pues así el haz es más ancho y se sigue con más facilidad.

Libros

Electroquímica, historia de la cámara oscura, teoría de campos y oceanografía regional

J. Obiols, J. L. Comellas, L. Mas y M. Alcaraz

Química Electroanalítica. Fundamentos y Aplicaciones, por P. Sánchez Batanero. Alhambra Universidad; Madrid, 1981. No es la electroquímica una materia fácil de enseñar. Concurren en ella diversos fenómenos y equilibrios químicos, cuya competencia en el sistema, aunque previsible desde un punto de vista termodinámico, no lo es en lo relativo a su cinética, por lo que resulta necesario disponer de datos experimentales para explicar los procesos reales. Mi experiencia en la enseñanza de la química analítica me confirma lo dicho: las lecciones dedicadas a los métodos electrométricos suelen ser donde los estudiantes hallan más arduo el camino, en particular las técnicas de trabajo, características del instrumental, previsión de interferencias y optimización de métodos. Dificultades atribuibles a la complejidad de los fenómenos que ocurren en la superficie de los electrodos y, en general, a la circulación de los electrones a través de la célula electrolítica. Valgan de ejemplo los fenómenos de sobretensión o de polarización.

Una de las formas clásicas de abordar, de un modo sistemático, el estudio de estos métodos es a través de las curvas intensidad-potencial (I/E), esquema que podríamos llamar "de la escuela francesa", iniciada por el profesor G. Charlot y seguida por numerosos científicos, entre los que está Sánchez Batanero, quien en esta obra realiza una presentación muy completa de esa línea didáctica: el estudio de los procesos electroquímicos a partir de la variación de la corriente que circula por un circuito potenciométrico a medida que varía el potencial aplicado sobre el electrodo de trabajo. Este esquema corre el riesgo de apoyarse demasiado en el cálculo gráfico, tópico en el que no incurre el autor, quien logra un buen equilibrio entre la presentación gráfica y la numérico-algebraica de los fenómenos estudiados.

Los cuatro primeros capítulos, de los doce en que se divide el libro, ocupan el 50 por ciento de la obra. Se centran en el estudio general de las curvas intensidad-potencial, para así entrar en la transferencia de masa y de carga en los electrodos; base teórica necesaria, no sólo para dominar las técnicas analíticas, sino también para profundizar en estudios de corrosión electroquímica al contemplar el comportamiento de los materiales en ingeniería química.

La segunda parte del libro versa acerca de los diferentes métodos electrométricos de análisis. Se establece una clasificación entre métodos indicadores y métodos cuantitativos; no se trata de una división entre métodos, sino diferentes aplicaciones que pueden darse a un mismo método. La exposición está orientada hacia los mecanismos electroquímicos en que se fundamentan los distintos procedimientos, sin desarrollar lo suficiente sus características analíticas, como la exactitud, la precisión y la sensibilidad; todo ello, junto con la falta de descripción de aplicaciones concretas, limita el libro al estudio teórico de las aplicaciones analíticas de la electroquímica. No se citan tablas de aplicaciones de las columbimetrías, amperometrías, etcétera, a las que podría acudir el lector para orientarse sobre problemas analíticos concretos. Es de esperar que en el volumen dedicado a problemas, cuya publicación se anuncia, se subsane esta laguna.

Se echa a faltar un capítulo dedicado a las conductimetrías, que si bien no están relacionadas con las curvas I/E, sí tienen su interés como método de análisis cuantitativo, en procesos industriales sobre todo. A pesar de que las curvas I/E son la base para una buena comprensión de las polarografías, esta técnica recibe un tratamiento superficial, omitiéndose sus variantes especiales que, como las voltametrías de "stripping" (eliminación), están mereciendo una destacada atención en la literatura analítica que se viene publicando en revistas especializadas. (J. O. S.)

The Camera Obscura. A Chronicle, Por John H. Hammond. Adam Hilger, Ltd. Bristol, 1981. 182 páginas. He aquí un libro sencillo, original, sin excesivas pretensiones, que no se cae de las manos. Escrito en un inglés limpio y cuidadoso, primorosamente editado, provisto de abundantes y valiosas ilustraciones, algunas de ellas nada fáciles de conseguir y en todo caso sumamente útiles para la comprensión del lector, versa sobre un tema de desarrollo aparentemente escaso, y de no gran actualidad, como es el de la cámara oscura.

El tratamiento que da a su estudio John Hammond –especialista en historia de la ciencia y del arte, y quizá sobre todo, en historia de instrumentos antiguos– es fundamentalmente histórico. De aquí la justificación del subtítulo, *A Chronicle*. Pero el libro de Hammond no se limita a una crónica de la cámara oscura, porque en este ingenio, como en la mayoría de las invenciones humanas, el orden lógico y el cronológico se superponen. Así, la obra puede concebirse también como una descripción de todos los modelos posibles de cámara oscura, desde los más sencillos y rudimentarios hasta los más sofisticados.

Obra de divulgación científico-cultural, puede interesar a toda persona curiosa, con independencia casi total de su dedicación especializada. Obviamente, despertará la atención de historiadores, historiadores de la ciencia, físicos, ópticos, fotógrafos, astrónomos, artistas, dibujantes, arquitectos, militares, museógrafos y coleccionistas. Este amplio espectro de lectores interesados en potencia puede expresar por sí solo el inesperado ámbito abarcado por el libro de Hammond; aunque esa misma amplitud, diríamos, lo hace singularmente apto para toda persona de cultura media o superior que desee un tipo de lectura instructiva, de viaje o vacaciones.

El autor divide su obra en nueve capítulos, dedicado el primero a los tiempos antiguos, y continuando luego, quizás un poco artificiosamente, por siglos, del XIII al XX, con un inciso –entre el XVII y el XVIII– dedicado a las aplicaciones de la cámara oscura al arte.

Los primeros en advertir la posibilidad de reproducir una imagen mediante el paso de la luz por un orificio fueron los pensadores chinos Chuen Mo-Ti, Tuang Cheng-Si o Yo Chao-Chang, en los que encontramos una mezcla curiosa de razonamientos lógicos, espíritu de observación e interpretaciones mágicas. La tendencia de los chinos a operar con luces y sombras –las "sombras chinescas"– les permitió ser los pioneros en la materia, sin llegar nunca a un verdadero planteamiento científico de la cámara oscura.

Aristóteles fue un observador más atento, y el primero en formularse preguntas. Intuyó claramente la simetría de los conos de luz cuyo vértice coincide en un estrecho orificio, y comprendió las razones de la formación de una imagen invertida sobre una pantalla colocada más allá del orificio en cuestión. Una experiencia que desazonó su espíritu clásico fue la de observar que, durante un eclipse de sol, los haces de luz filtrados por una "cámara oscura" tan irregular como las hojas de un árbol proyectaban sobre el suelo una imagen correcta de la fase solar. El motivo de tan inesperada perfección no pudo ser explicado hasta fines del siglo XVII.

En la décima centuria, el físico y matemático árabe Alhazen, operando con tres focos de luz simultáneamente, observó los resultados de su proyección a través de un orificio o ranura, e intuyó por primera vez la estricta linealidad de la luz. Sin embargo, fueron los occidentales, de Roger Bacon a Leonardo de Vinci –pasando por Alberti y Vitelio–, quienes dieron forma y fundamento a la cámara oscura propiamente dicha, explicando cada vez mejor su teoría, al tiempo que perfeccionaban su fabricación. A partir del siglo XVI, la cámara fue ya un instrumento habitual al servicio del hombre, utilizado lo mismo como curiosidad y recreo que como elemento científico –para la observación y correcta reproducción de imágenes, especialmente eclipses– o artístico. Fabricius fue el primero en observar las manchas solares por medio de la cámara oscura, y Gianbattista della Porta llevó a tal perfección y difundió tan ampliamente el ingenio, que pasa –erróneamente– por ser su descubridor.

La aplicación de lentes y espejos permitió a la cámara oscura unas posibilidades hasta entonces desconocidas. En el siglo XVII se construyen ya cámaras portátiles y se penetra a fondo en sus fundamentos ópticos. Jeremías Horrox presenció en una cámara oscura el paso de Venus ante el disco del Sol, el 24 de noviembre de 1639; y R. Hooke fue el primero en descubrir las ventajas de una pantalla cóncava. También de la misma época data la concepción del ojo humano como una cámara oscura, y de la retina como pantalla. Hombres tan destacados como Kepler o Descartes no dejaron de experimentar con el ya difundido instrumento; y parece que fue Kepler –observador de un paso de Mercurio– el introductor del término "cámara oscura".

Constituye un grato intermedio el capítulo dedicado a las aplicaciones de la cámara al arte, especialmente a la pintura, y en menor grado a la arquitectura. Hammond estudia no sólo el empleo de cámaras de proyección, sino otros ingenios basados en fundamentos ópticos distintos, para apreciar correctamente la perspectiva, o para obtener reproducciones monoculares a escala exacta. Puede ser para muchos lectores una curiosa sorpresa saber que artistas como Vermeer o Canaletto pintaban con frecuencia no directamente el objeto, sino su imagen vista a través de una cámara oscura. El recurso puede descubrirse, por ejemplo, en las inversiones derecha-izquierda de los paisajes reproducidos. Y de aquí deriva no sólo la precisión lineal de la obra de estos pintores, sino también la luminosidad o la especial delicadeza de sus lienzos.

El siglo XVIII señala la edad de oro de la cámara oscura, con la fabricación y generalización de instrumentos de todo tipo, algunos particularmente ingeniosos. Se hizo frecuente que las instituciones científicas poseyeran una gran cámara oscura, a la que acudían grupos de visitantes, con una curiosidad similar a la que hoy suscitan, por ejemplo, los planetarios. No faltan viajeros que llevan cámaras oscuras portátiles a los más diversos países, y, al tiempo que las utilizan para reproducir imágenes de paisajes o monumentos, hacen un saneado negocio como organizadores de pequeños espectáculos públicos.

Esta tendencia se generaliza en el siglo XIX, cuando no hay feria o exposición que se precie si no dispone de una gran cámara oscura, donde el público puede contemplar admirado, sobre una pantalla, el mismo paisaje que acaba de ver fuera con sus ojos. La programación de imágenes móviles –mucho más efectistas, y fuente a veces de reacciones histéricas por parte de los espectadores– puede ser el primer precedente de las técnicas cinematográficas. Al mismo tiempo, Daguerre y Talbot, al colocar una emulsión sensible sobre la pantalla, dan a la cámara oscura su más universal aplicación contemporánea. con el nacimiento de la fotografía.

No es este, sin embargo, el único servicio de la cámara oscura en los siglos XIX y XX. Hammond nos describe los usos militares de que ha sido objeto, desde las observaciones detalladas del funcionamiento de artefactos explosivos desde un recinto protegido, hasta el registro del movimiento de navíos desde observatorios invisibles, u otros servicios de espionaje, frecuentes en la primera guerra mundial.

Hoy la cámara oscura propiamente dicha ha perdido actualidad, y son pocas las que, con una finalidad científica, o al servicio de la curiosidad pública, se conservan operativas en el mundo. Muchas de ellas han pasado a ser objetos de museo. Por el contrario, las aplicaciones de su principio han alcanzado su máximo desarrollo en nuevas formas de instrumental, entre el que debemos contar diversos tipos de máquinas fotocopiadoras.

Si algo puede objetarse a la obra de Hammond es un cierto linealismo descriptivo, que nos dificulta la obtención, para un momento dado, de una visión panorámica del tema. Por motivos tal vez justificados, pasa como sobre ascuas por dos temas que implican un sustancial cambio de naturaleza en la cámara oscura: la introducción de una lente sustituyendo al simple orificio y el descubrimiento de la fotografía. Cabría, efectivamente, distinguir dos tipos ópticamente muy diferenciados de cámara, según que la imagen esté formada por simple proyección, o por refracción; pero uno de los puntos que el autor soslaya sistemáticamente es justo el concepto o definición de "cámara oscura": omisión que le permite referirse con plena libertad a pínulas, o hasta a las imágenes formadas en el suelo por la sombra del follaje. Por lo que se refiere a la adaptación de la cámara a la técnica fotográfica, al autor apenas le interesa otro extremo que dejar en claro el papel de Talbot, a su juicio más completo que el de Daguerre, en su invención.

Un elenco de notas, reducido, pero suficiente, y una detallada bibliografía avalan la buena información del autor y enriquecen el valor científico de su trabajo. (J. L. C.)

LAS TEORÍAS DE LOS CAMPOS DE FUERZA, DESDE FARADAY HASTA EINSTEIN, por William Berkson. Alianza Universidad; Madrid, 1981. Dentro de una rama de la ciencia podemos clasificar a los libros que sobre ella versan en libros de investigación, libros de texto a diferentes niveles, libros de historia de esa ciencia y libros de filosofía sobre la

misma, que, a su vez, puede dividirse entre filosofía de las teorías antiguas y filosofía de las teorías actuales. El libro que nos ocupa tiene un poco de todo, pero más de historia y filosofía de una cierta rama de la física, la teoría de campos, tal como se la enunciaba en el siglo XIX y principios del nuestro. Insiste el autor en que los libros de texto deberían incluir los datos y razonamientos históricos; en algún punto aprovecha la ocasión para introducir teorías propias sobre el problema del éter o ideas metafísicas personales.

A través de la clasificación de las observaciones, la física intenta obtener modelos o teorías que expliquen el mundo desde un punto de vista científico. A lo largo de la historia ha habido bastantes teorías y no todas defendían la misma idea de universo. En el ámbito de la física moderna, que incluiría las teorías cuánticas y las teorías relativistas, abundan las obras a todos los niveles citados, desde una divulgación elemental hasta libros de texto para licenciados, pasando por los numerosos artículos que sobre física moderna se publican en revistas como *Investigación y Ciencia*. Puede que en 1974, fecha de edición del original del libro que comentamos, no hubiera tantos, según reconoce el autor.

Que toda teoría sea y deba ser cuestionada constituye un principio que la comunidad científica acepta y utiliza. Pero no considero correcto que se continúe aludiendo a Einstein, Schrödinger o de Broglie como únicos impugnadores de la teoría cuántica. Muchos de sus planteamientos han sido superados y otros nuevos han aparecido o evolucionado apreciablemente. También creo que la mayoría de los físicos aceptan hoy como bastante coherentes las teorías relativistas, la relatividad especial y la relatividad general, y las teorías cuánticas, que incluirían todas las relaciones con cuantos.

La relatividad especial también sigue cuestionada, no sólo en un sentido negativo (el de pretender suprimirla), sino también en el sentido positivo, de intentar mejorarla. En ello entra lo que de teoría de campos tenga la teoría de la relatividad; baste recordar los nombres de Feynman, Wheeler y Wigner, que iniciaron un camino de vuelta hacia la acción a distancia compatible con la relatividad especial y compatible también con un determinismo que, casi sin darse cuenta sus creadores, había perdido la física *clásica* (no cuántica) al introducir la teoría de campos. Considérese sino el problema de dos cuerpos, solos en el universo, y que se mueve cada uno dentro del campo creado por el otro y puede que dentro del autocampo. Sólo si imponemos como hipótesis que los cuerpos se muevan en una sola dimensión (hay que añadir, pues, la condición de que son cuerpos puntuales sin dimensiones), en algunos casos está demostrado el buen planteamiento, es decir, que las ecuaciones que se deducen de la correspondiente teoría de campos admiten solución [Luis Bel: *Ann. Inst. H. Poincaré, XII*, 307, 1970].

Estas dificultades no objetan que la teoría de campos haya sido superada, siempre que incluyamos ahí las varias teorías y los varios campos actualmente considerados, sin soslayar la métrica como campo del espacio-tiempo en la teoría de la relatividad general. Debemos señalar con todo que, para dirimir entre campo y acción a distancia, no es crucial el concepto de velocidad de propagación instantánea o no, tal como indica Berkson en la introducción y corrige bastante más adelante, ya que existen teorías de acción a distancia con un tiempo de propagación no nulo de la acción.

Importa señalar que "al margen de las aplicaciones tecnológicas, el mayor atractivo de la ciencia reside en el descubrimiento de aspectos desconocidos del mundo". Lo que hoy tiende a olvidarse, exigiéndosele a ciertas ramas del saber una gran rentabilidad inmediata. Pretensión no exenta de peligro, si nos fijamos en quién va a utilizar dicha rentabilidad, y que resta valor al desarrollo intrínseco de la ciencia: aquel que aporta un mejor conocimiento de la realidad que se nos pone delante.

Por su actualidad, vale la pena recordar la idea que Meyerson tenía sobre la explicación científica. A partir de unos resultados experimentales se empieza a construir un modelo asociando las magnitudes físicas a determinados conceptos matemáticos; con las hipótesis del modelo o teoría se hacen deducciones matemáticas que luego se restituyen a las magnitudes físicas que han originado el estudio, de cara a la comprobación experimental del modelo introducido.

Puede haber dificultades en el paso del nivel matemático del modelo al nivel físico o real por dos razones: (*a*) Si el modelo o teoría es muy general, ofrece muchos caminos hacia el mundo físico, lo que dificulta la comprobación experimental; la disconformidad de los resultados en los dos niveles citados puede deberse a un fallo en la elección de la correspondencia adecuada; ejemplo claro de esta dificultad lo presenta la teoría de la relatividad general. (*b*) La asidua frecuencia con que cargamos cierta magnitud o concepto matemático de un contenido que no tiene y, al pasar al nivel físico, le exigimos un comportamiento indebido según las hipótesis de la teoría; pensemos en las nociones de tiempo, espacio y otros que, empleadas por personas ajenas a la teoría, agravan la dificultad mencionada. Es lo que Berkson llama la confusión entre ciencia y metafísica. La teoría cuántica actual ofrece la ventaja de eludir casi completamente ese último inconveniente, ya que utiliza conceptos matemáticos muy alejados de los que se podrían parecer a ciertas magnitudes físicas corrientes. Ello se traduce también en un número de detractores menor, comparado con los que tiene la teoría de la relatividad especial.

Intentar explicar las muchas ideas que se le ocurrían a Faraday es tarea ardua. El autor consigue en algunos casos introducirnos en la compleja, intuitiva y nada matemática mente de Faraday; no así en otros. Es empresa difícil condensar en pocas páginas una línea de trabajo, y de trabajo intenso como el que realizó Faraday. Ni es cómodo explicar líneas de trabajo que luego no fructificaron. En efecto, el propio inventor del camino, aquí Faraday, a veces se da cuenta de que anda errado, y por tanto el historiador, aquí Berkson, en aras de una mejor comprensión, debe aclarar o terminar lo que ni el mismo inventor resolvió. Además, le cuesta mucho al lector poner su formación entre paréntesis, de suerte que sus estudios actuales de la teoría no interfieran, y situarse en el contexto de la época en que se estaban planteando las consideraciones que se explican.

Nunca abordó Faraday en sus escritos, de una forma rigurosa, las ideas metafísicas que podrían deducirse de sus resultados. No nos ha llegado ninguna de las muchas ideas que debieron acudírsele sobre la teoría de campos, lo que induce a pensar que quizá Faraday no compartiera el entusiasmo del autor por el aspecto instructivo que caracterizase al libro de texto donde se recogiera la metafísica asociada a sus resultados experimentales y a sus ideas. Pondérese también cuán difícil resulta exponer y polemizar sobre los conceptos metafísicos de Faraday, cuando no nos dejó por escrito lo que bullía en su cabeza.

De la historia social destacaría el interés que había en la Inglaterra del XIX por la educación de la clase obrera, plasmado en la organización de sesiones muy bien preparadas y que dieron

sus frutos. Cabría preguntarse cómo está ahora este asunto, y no sólo en Inglaterra. Considero otro acierto señalar las ideas que sobre masa y fuerza introdujo Boskovitch, bastante olvidadas en un nivel de introducción a la historia de la física, y que incluyen, entre otras, la noción entonces inaudita de la existencia posible de varias partículas en un mismo lugar, partículas que luego llamaríamos bosones. Se sigue investigando dentro de esta línea, al objeto de superar la dificultad del tamaño puntual de las partículas elementales sin trascender el plano completamente clásico, no cuántico, buscando modelos clásicos de partículas elementales [A. F. Rañada: *J. Math. Phys. 22,* 2533, 1981].

Cuando en la introducción el autor habla del determinismo de la mecánica de Newton debemos denunciar un grave error en el enunciado y explicación de la segunda ley de Newton (que él indica como primera). Las condiciones iniciales que fijan la evolución no son la fuerza y la masa, que tienen que ser conocidas *en todo instante*, sino la posición y velocidad de la partícula en un instante inicial. Cuando mantiene que, en todo rigor, no hay contradicción en la utilización del concepto de línea de fuerza, tampoco estoy de acuerdo ni en la expresión de *en todo rigor*, ya que una contradicción o es o no es, ni en que se ignore que el concepto de campo electromagnético es el único intrínseco, perdiendo este carácter que podían haber tenido antes los campos eléctrico y magnético por separado; en cambio sigue teniendo su vigor conceptual la línea de corriente de un fluido.

En este mismo sentido, discrepo de su afirmación según la cual la teoría de la relatividad especial incluye como gran novedad la pérdida del carácter unívoco de ciertos conceptos como masa y volumen. Berkson señala, después, que conocida la masa y volumen de un cuerpo en un determinado sistema de referencia, aunque cambien, sabemos calcular su valor en otro sistema de referencia; pero es que además hay un sistema de referencia especial en el que los valores que se miden tienen un sentido especial, que es el sistema en el que el cuerpo está quieto; en dicho sistema los valores medidos reciben el calificativo de *propios*, y, en general, cuando se habla de la masa o del volumen de un cuerpo, aunque no se añada el calificativo de propios se sobreentiende, por lo que siempre estamos aludiendo a las mismas cantidades.

La simpatía con que el autor retrata el comportamiento de Heaviside, en abierto contraste con el de algunos de sus contemporáneos que demostraban en sus artículos una falta de coherencia y una cantidad de errores considerable, se convierte luego en comprensión hacia los detractores de la teoría de la relatividad especial, que en la mayoría de los casos presentan argumentos totalmente acientíficos, no siendo despreciable el número de los que en su afán por corregir o incluso destruir la teoría avanzan argumentos que contradicen la dinámica de Newton en la zona de solapamiento de ambas teorías. Hay que vigilar y estudiar estas zonas de solapamiento, pero esta vigilancia y estudio incluyen ya, como nueva zona para una teoría más reciente, a las dos o más anteriores. En un libro histórico, tanto de la ciencia como de la filosofía de la ciencia, sobran las ideas personales que uno pueda tener acerca de las nuevas teorías. Sin necesidad de seguir una línea positivista a ultranza puede uno discrepar profundamente de lo que expone Berkson sobre una teoría del éter móvil cuya velocidad es *en principio indetectable*.

No creo que un libro exclusivamente histórico o de la filosofía de la ciencia estimule ideas nuevas de la física actual; basta fijarse en cómo la física contemporánea ha obtenido la mayoría de sus resultados. El asentamiento y clarificación de nuevas teorías, tema ineludible, requiere, no sólo una buena formación histórico-filosófica, sino también un conocimiento cabal de los resultados científicos. Un estudio a posteriori de toda nueva teoría siempre encuentra un camino histórico que la justifica. Y si bien, como dice el autor, "los problemas antiguos restringen las posibles soluciones de los nuevos", ello sólo se refiere a algunos que a priori no podemos determinar. A veces sucede todo lo contrario.

En la mayoría de libros de texto se introducen notas históricas, algunas erróneas como la que Berkson indica sobre el libro de electrodinámica clásica de Jackson, pero en éste caso hay que señalar que el error de Jackson no ha influido demasiado en los estudiantes que habiéndolo utilizado luego obtuvieron buenos resultados en sus investigaciones, y que no fueron pocos. El concepto de metafísica del autor es algo restringido, aunque adecuado a la época que describe; hoy en día se admite que cada ciencia requiere modelos distintos o difícilmente superponibles. Dentro de la física parece cada vez más difícil hablar de una sola rama y, aunque en partículas elementales se tienda a una teoría unificada, no creo que ésta sea aplicable a los fenómenos biofísicos, por citar un ejemplo.

El autor pretende que su obra debiera ser llana para quienes estén familiarizados con las ecuaciones de Maxwell. Yo dudo que ni siquiera para éstos sea así. Téngase en cuenta que en su brevedad se condensa una cantidad ingente de ideas, cuya ordenación y síntesis siempre resulta arriesgada. El libro es polémico en bastantes puntos, lo que es muy bueno, pero la dificultad de la lectura le resta mérito. (L. M.)

OCEANOGRAFÍA DEL MAR MEDITERRÁNEO, por Jaime Rodríguez. Ediciones Pirámide, S.A. Madrid, 1982. 174 páginas. El interés por el conocimiento científico del medio marino, que tuvo su primera gran manifestación en el último cuarto del siglo pasado con la expedición del *Challenger* (1872-1876), coincidió casi simultáneamente con la creación de centros o estaciones costeras permanentes, dedicadas a la investigación de la biología marina y oceanografía. El primer instituto se estableció precisamente en el litoral mediterráneo (la Stazione Zoologica di Napoli, fundada en 1872).

Desde aquella época, de intensa creatividad, en la que las ciencias del mar recibieron un impulso considerable (al que tal vez no fuese ajeno el "touch of class" que le conferían algunas testas coronadas europeas interesadas en el tema), se ha pasado al esfuerzo colectivo de organizaciones internacionales y a la existencia de casi ciento cincuenta instituciones, repartidas por todo el Mediterráneo, que han contribuido a crear un sólido cuerpo de conocimientos, aunque éstos apenas se han concentrado en trabajos de síntesis.

El libro que nos ocupa, *Oceanografía del mar Mediterráneo*, como indica el autor en el prólogo, es una consecuencia directa de su labor docente (es profesor adjunto de ecología de la Universidad de Málaga), lo que se refleja en parte en la estructura de la obra; en ella pueden distinguirse tres grupos de temas, que corresponden respectivamente al medio físico, a la biología y al impacto del hombre sobre el *mare nostrum.* El libro está dividido en seis capítulos, con la bibliografía correspondiente al final de cada uno de ellos, más un glosario de términos científicos utilizados, que cierra la obra.

De los capítulos dedicados a la descripción del medio físico, el primero trata del origen, morfología y estructu-

ra sedimentaria de las diferentes cuencas que forman el Mediterráneo, así como de la climatología de las diferentes áreas, mientras que en el segundo se ofrece una panorámica de la hidrografía (relaciones con el océano Atlántico, origen, circulación, formación de masas de agua mediterráneas, balance de sales, etcétera). Los aspectos biológicos se abordan en los capítulos tres, cuatro y cinco; en el tercero se compara la fertilidad potencial del Mediterráneo con la de otras áreas marinas y se discute el origen de las sales nutritivas, así como los mecanismos de fertilización, terminando con una breve exposición de las especiales características de la estructura y función de las poblaciones de fitoplancton y oportunos comentarios sobre las "mareas rojas" o "purgas de mar" y las condiciones que favorecen su aparición. Se compara la fauna marina mediterránea con la de áreas marinas vecinas desde el punto de vista biogeográfico (afinidades y posible origen) en el capítulo cuarto, y en el quinto se describe, de forma un tanto rápida, la composición y estructura de los dos grandes sistemas marinos, el pelágico y el bentónico.

En el capítulo seis, después de una prudente declaración de principios en la que el autor aclara su postura respecto del binomio ecología-ecologismo, se incluye al hombre como un elemento más del ecosistema marino. En este capítulo se examinan tanto los efectos de la explotación humana de los recursos renovables como la situación actual del Mediterráneo en lo concerniente a su contaminación. El glosario final contiene unos sesenta términos científicos, los de carácter más específico entre los utilizados en el libro, convenientemente explicados.

Debe destacarse el esfuerzo considerable realizado por el autor, tanto por la diversidad de los problemas abordados como por las dificultades que entraña el resumir temas tan vastos en un corto, tal vez demasiado corto, número de páginas. Resulta muy somero el tratamiento que reciben algunos temas, el del zooplancton por ejemplo, precisamente la especialidad del autor, quien parece haber sacrificado, quizás en aras de la brevedad, algunas explicaciones de índole general que tal vez hubiesen facilitado el acceso al libro de un público no familiarizado con la ecología. Algunos conceptos se sintetizan casi con violencia, lo que no impide que consiga una valiosa panorámica de los conocimientos actuales sobre la oceanografía mediterránea. (M. A.)

Bibliografía

Los lectores interesados en una mayor profundización de los temas expuestos pueden consultar los trabajos siguientes:

EL TELESCOPIO ESPACIAL

SCIENTIFIC RESEARCH WITH THE SPACE TELESCOPE: IAU COLLOQUIUM No. 54. Dirigido por M. S. Longair y J. Warner. National Aeronautics and Space Administration CP-2111, 1979

THE SPACE TELESCOPE SCIENTIFIC INSTRUMENTS. David S. Leckrone en *Publications of the Astronomical Society of the Pacific*, vol. 92, n.° 545, págs. 5-21; febrero-marzo, 1980.

SUPERCONDUCTORES ORGANICOS

THE DESIGN AND SYNTHESIS OF ORGANIC METALS. Anthony F. Garito y Alan J. Heeger en *Accounts of Chemical Research,* vol. 7, n.° 7, págs. 232-240; julio, 1974.

THE DIFFERENCE BETWEEN METALLIC AND INSULATING SALTS OF TETRACYANOQUINODIMETHANE (TCNQ). Jerry B. Torrance en *Accounts of Chemical Research,* vol. 12, n.° 3, págs. 79-86; marzo, 1979.

THE PHYSICS AND CHEMISTRY OF LOW DIMENSIONAL SOLIDS: PROCEEDINGS OF THE NATO ADVANCED STUDY INSTITUTES HELD AT TOMAR, PORTUGAL, AUGUST 26-SEPTEMBER 7, 1979. Dirigido por Luis Alcacer. D. Reidel Publishing Co., 1980.

EXTENDED LINEAR CHAIN COMPOUNDS. Joel S. Miller. Plenum Press, 1982.

AVIONES ULTRALIGEROS

PROGRESS IN FLYING MACHINES. Octave Chanute. *The American Engineer and Railroad Journal*, 1894. Reimpresión facilitada por Ultralight Publications, P.O. Box 234, Hummelstown, Pennsylvania 17036.

ULTRALIGHT AIRCRAFT: THE BASIC HANDBOOK OF ULTRALIGHT AVIATION. Michael A. Markowski. Ultralight Publications, P.O. Box 234, Hummelstown, Pennsylvania 17036, 1981.

EFUSIONES DE ENERGIA EN ESTRELLAS JOVENES

LA GALAXIA DE LA VIA LACTEA. Bart J. Bok en *Investigación y Ciencia*, n.° 56, págs. 54-72; mayo, 1981.

COMPLEJOS GIGANTES DE NUBES MOLECULARES EN LA GALAXIA. Leo Blitz en *Investigación y Ciencia*, n.° 69, págs. 48-58; junio, 1982.

TOPOISOMERASAS DE ADN

STRUCTURE AND FUNCTIONS OF DNA y BINDING AND UNWINDING PROTEINS AND TOPOISOMERASES. Arthur Kornberg en *DNA Replication.* W. H. Freeman and Company, 1980.

ADN SUPERENROLLADO. William R. Bauer, F. H. C. Crik y James H. White en *Investigación y Ciencia*, n.° 48, págs. 78-89; septiembre, 1980.

DNA TOPOISOMERASES. Martin Gellert en *Annual Review of Biochemistry,* vol. 50, págs. 879-910; 1981.

LA SECRECION INTESTINAL

INTESTINAL SECRETION. T. R. Hendrix y H. T. Paulk en *International Review of Physiology. Gastrointestinal Physiology II*, vol. 12, dirigido por R. K. Crane. University Park Press; Baltimore, 1977.

MECHANISM OF SELECTIVELY INHIBITING THE ACTIVATION OF CYCLIC NUCLEOTIDE PHOSPHODIESTERASE AND ADENYLATE CYCLASE BY ANTIPSYCHOTIC AGENTS. B. Weiss y R. M. Levin en *Advances in Cyclic Nucleotide Research*, vol. 9, págs. 285-303, dirigido por W. J. George y L. J. Ignarro. Raven Press, Nueva York, 1978.

ROLE OF CA^{2}-DEPENDENT REGULATOR PROTEIN IN INTESTINAL SECRETION. A. Ilundáin y R. J. Naftalin en *Nature*, vol. 279, págs. 446-448; 1979.

CALMODULIN PLAYS A PIVOTAL ROLE IN CELLULAR REGULATION. W. Y. Cheung en *Science*, vol. 207, págs. 19-27; 1980.

LA REPRODUCCION COOPERATIVA DE LA ABUBILLA ARBOREA VERDE

THE COMMUNAL SOCIAL SYSTEM OF THE GREEN WOODHOOPOE IN KENYA. J. David Ligon y Sandra H. Ligon en *Living Bird*, vol. 16, págs. 159-197; 1978.

DEMOGRAPHIC PATTERNS AND COMMUNAL BREEDING IN THE GREEN WOODHOOPOE, *PHOENICULUS PURPUREUS*. J. D. Ligon en *Natural Selection and Social Behavior: Recent Research and New Theory*, dirigido por R. D. Alexander y D. W. Tinkle. Chiron Press, Inc., 1981.

ECOLOGICAL FACTOR AND KIN SELECTION IN THE EVOLUTION OF COOPERATIVE BREEDING IN BIRDS. W. D. Koenig y F. A. Pitelka en *Natural Selection and Social Behavior: Recent Research and New Theory*, dirigido por R. D. Alexander y D. W. Tinkle. Chiron Press, Inc., 1981.

ANEURISMAS

CONTRIBUTION OF ABDOMINAL AORTIC ANEURYSMECTOMY TO PROLONGATION OF LIFE. D. Emerick Szilagyi, Roger F. Smith, Franklin J. DeRusso, Joseph P. Elliot y Frederick W. Sherrin en *Annals of Surgery*, vol. 164, n.° 4, págs. 678-699; octubre, 1966.

ANEURYSMS OF THE AORTA. Crile Crisler y Henry T. Bahnson en *Current Problems in Surgery*, diciembre, 1972.

SURGICAL MANAGEMENT OF ABDOMINAL AORTIC ANEURYSMS: FACTORS INFLUENCING MORTALITY AND MORBIDITY— A 20-YEAR EXPERIENCE. Jesse E. Thompson, Larry H. Hollier, R. Don Patman y Alfred V. Persson en *Annals of Surgery*, vol. 181, n.° 5, págs. 654-661; mayo, 1975.

INFRARENAL AORTIC ANEURYSMS. Robert B. Rutherford en *Vascular Surgery*, dirigido por Robert B. Rutherford. W. B. Saunders Company, 1977.

ORIGIN, GROWTH AND RUPTURE OF SACCULAR ANEURYSMS: A REVIEW. Laligan N. Sekhar y Roberto C. Heros en *Neurosurgery*, vol. 8, n.° 2, págs. 248-260; febrero, 1981.

TEMAS METAMAGICOS

HANDBOOK THE CUBIK MATH. Alexander H. Frey, Jr., y David Singmaster. Enslow Publishers, 1982.

INSIDE RUBIK'S CUBE AND BEYOND. Christoph Bandelow. Birkhauser Boston, Inc., 1982.

PUZZLE IT OUT: CUBES, GROUPS AND PUZZLES. John Ewing y Czes Kośniowski. Cambridge University Press, 1982.

TALLER Y LABORATORIO

BOUNCING LIGHT BEAM. William M. Strouse en *American Journal of Physics*, vol. 40, n.° 6, págs. 913-914; junio, 1972.

TRIBOLUMINESCENCE SPECTROSCOPY OF COMMON CANDIES. Rebecca Angelos, Jeffrey I. Zink y Gordon E. Hardy en *Journal of Chemical Education*, vol. 56, n.° 6, págs. 413-414; junio, 1979.